AN EXAMINATION OF EVALUATION PLAN QUALITY IN HUMAN SERVICE
NONPROFIT ORGANIZATIONS

Sarah M. Harrison

A Dissertation
Submitted to the Graduate College of Bowling Green
State University in partial fulfillment of
the requirements of the degree of

DOCTOR OF EDUCATION
May 2009

Committee:
Judith A. Zimmerman, Advisor
Catherine Cassara-Jemai
Graduate Faculty Representative
Sarah Twitchell
Rachel A. Vannatta
ABSTRACT

Judith A. Zimmerman, Advisor

The purpose of this correlational study was to examine the relationship between organization size, size of funding allocation, type of human service organization, organization staff position responsible for the evaluation plan and evaluation plan quality in human service nonprofit organizations in Ohio. Evaluation plan quality was assessed by using the Program Accountability Quality Scale (PAQS). This study utilized archival data from 125 human service nonprofit organizations collected by a United Way organization to examine factors related to evaluation plan quality.

The research findings demonstrated that evaluation plan quality did not relate to organization size. Therefore, although there were differences in relation to financial resources between the organizations it did not impact evaluation plan quality. In regard to the relationship between size of funding allocation and evaluation plan quality, the activities and outputs subscales resulted in lower quality scores for organizations that received smaller funding allocations when compared to those organizations that received the largest funding allocations. This study demonstrated some significant differences between type of human service organization and evaluation plan quality. While this study contributed to the literature regarding the role and use of evaluation efforts in the nonprofit sector, future research is warranted.

The following conclusions were drawn from this study that examined the relationship between evaluation plan quality and organization factors: 1) the size of the human service organization was not related to evaluation plan quality, 2) while the results of the relationship between funding allocation and evaluation plan quality varied, it did appear that on a few of the measures organizations that received less funding exhibited lower quality evaluation plans, 3)
there were some significant differences between some types of human service organizations and evaluation plan quality and 4) there is a critical need for further research to understand evaluation systems utilized by nonprofit organizations.
I dedicate this dissertation to my spouse, Ruben Viramontez Anguiano.

His support and encouragement made this possible.
ACKNOWLEDGMENTS

Thanks to my dissertation committee, Drs. Judy Zimmerman, Rachel Vannatta, Catherine Cassara-Jemai and Sarah Twitchell that offered guidance, support and asked challenging questions. Although my passion for the field of nonprofit organizations was different from their interests, my committee provided critical insight that helped form this research study.

My cohort members were a source of motivation and inspiration throughout this process. A special thank you for your friendship and support during the many Tuesday nights that we spent together. Each of the cohort members, in their own unique way, provided encouragement that helped me reach the end of this journey.

A special thank you to those members of the nonprofit community that provided guidance and ultimately made this study a reality. Thanks to those that provided advice along the way, particularly Karis Box, Dr. Randy Stoecker, Joan Nelson and many others. And I am extremely grateful for the staff at the United Way that were willing to share data and information they had collected. Their openness and generosity is to be commended and speaks to the impact they make in the community. Also to the Heart of Florida United Way that granted permission to use the Program Accountability Quality Scale instrument.

Many friends and family members offered support along the way. A big thank you to all of you that asked about my progress during this challenging journey, your encouragement provided much needed motivation and support. A special thank you to my parents, Dan and Josie Harrison for always believing in me. I will forever be grateful for the support and sacrifices that were made by my family: to Ruben for creating the space in our lives for me to complete this degree and to my sons Lorenzo and Olin for reminding me about the important things in life.
# TABLE OF CONTENTS

**CHAPTER I. INTRODUCTION**

<table>
<thead>
<tr>
<th>Background of the Problem</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rationale for the Study</td>
<td>3</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>4</td>
</tr>
<tr>
<td>Research Questions</td>
<td>4</td>
</tr>
<tr>
<td>Theoretical Framework</td>
<td>4</td>
</tr>
<tr>
<td>Significance of the Study</td>
<td>5</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>6</td>
</tr>
<tr>
<td>Delimitations</td>
<td>7</td>
</tr>
<tr>
<td>Limitations</td>
<td>8</td>
</tr>
</tbody>
</table>

**CHAPTER II. LITERATURE REVIEW**

| Understanding the Nonprofit Sector | 9 |
| Demand for Accountability and Evaluation | 10 |
| Influence of Business Evaluation on the Nonprofit Sector | 12 |
| Relationship between Nonprofit Organizations and Funders | 15 |
| Use of Evaluation Systems         | 17 |
| Development of Program Based Human Service Evaluation Systems | 18 |
| Concerns with Nonprofit Evaluation | 20 |
| The Use of Evaluability Assessment | 20 |
| The Role of Stakeholders          | 21 |
| Usefulness of Evaluation and Data Collection Efforts | 23 |
**LIST OF TABLES**

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Descriptive Statistics of Human Services Organizations and Evaluation Plan</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>Quality</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Correlation Coefficient of Organization Size and Evaluation Plan Quality</td>
<td>48</td>
</tr>
<tr>
<td>3</td>
<td>Analysis of Variance Examining Significant Differences Between Organization</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>Size and Evaluation Plan Quality</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Correlation Coefficient of Funding Allocation and Evaluation Plan Quality</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>Analysis of Variance Examining Significant Differences Between Funding</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Allocation Groups and Evaluation Plan Quality</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Analysis of Variance Examining Significant Differences Between Type of Human</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>Service Organization and Evaluation Plan Quality</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER I. INTRODUCTION

Background of the Problem

In the United States, nonprofit organizations have a distinct role in society as they have been created to benefit the public or their members through a range of services (Czerwinski, 2007). Human service nonprofit organizations are of particular importance as they provide services to society’s most vulnerable citizens. Human service nonprofit organizations typically provide programs, services and support intended to meet the physical and psychological needs of individuals. These types of organizations often have a wide range of missions and include organizations such as homeless shelters, food banks, after school enrichment programs, mental health counseling and home health care ("Facts and figures about charitable organizations", 2007; Plantz, Greenway, & Hendricks, 1997).

In the last two decades, human service organizations have increasingly been expected to determine the impact of their services in a quantifiable way. Funding sources, individual donors and the public have requested and at times required that organizations move beyond a financial reporting of expenditures to adopt a more complex evaluation system that encompasses both financial and programmatic aspects (Cutt & Murray, 2000; Poole, Davis, Reisman, & Nelson, 2001; Thayer & Fine, 2001). Human service organizations have attempted to meet this demand for accountability through the use of a variety of program based evaluation systems (Bozzo, 2000; Campbell, 2002; Stoecker, 2007).

While evaluation has become more commonplace in the nonprofit sector, organization leaders have been hesitant to adopt evaluative measures, because of concerns that performance results are based on unrealistic standards, yet will be used as a primary determinate in funding allocations (Campbell, 2002; Cutt & Murray, 2000; Julian & Kombarakaran, 2006; Wholey &
Newcomer, 1997). Nonprofit leaders state that while decreasing or eliminating funding to lackluster programs is understandable, there are concerns that the evaluation results do not adequately represent a program’s impact (Wholey & Newcomer, 1997). Without proven methods to evaluate the quality of evaluation systems and a clear understanding of the purpose of the evaluation, organizations are likely to develop evaluation measures referred to as “goal displacement” evaluation. Goal displacement evaluation occurs when the goals of evaluation are adjusted in anticipation of the intended use of the results. For example, in goal displacement evaluation measures are designed with the intention of demonstrating success rather than attempting to measure the impact of the program (Cutt & Murray, 2000; Newcomer, 1997; Wholey & Newcomer, 1997). Due to increasing requirements to engage in and report evaluation results, it is imperative that the nonprofit sector seeks to implement quality evaluations.

When appropriate, quality evaluations occur within a human service organization, the results can be utilized in a variety of ways. Evaluation results can be used by stakeholders, including funding entities and individual donors, to inform future financial contribution decisions. Financial support from stakeholders can have a direct impact on sustaining, eliminating or expanding a program. Evaluation results can also be used by organization leaders to inform strategic decision making and capacity building initiatives. Recognizing the multifaceted impact that evaluation results potentially may have on the services and programs offered by human service organizations, nonprofit sector leaders have drawn attention to the importance of examining the quality of evaluation systems utilized by organizations. Without an analysis of evaluation quality, it is inappropriate to make critical funding decisions about human service programs that impact the lives and well-being of individuals and families. While current research highlights the importance of evaluation quality, little research has been completed to
examine organizational factors that impact evaluation plan quality. Evaluation plan quality has been defined as the feasibility, usefulness and validity of a system designed to measure program services (Poole, Nelson, Carnahan, Chepenik, & Tubiak, 2000). Recognizing the importance of having a quality evaluation plan prior to data collection and analysis, it is critical that in depth study of evaluation planning occur. This study was intended to further the limited research that exists in the area of evaluation plan quality and explored factors related to evaluation plan quality in human service nonprofit organizations.

Rationale for the Study

This study provided insight into factors that were related to the quality of evaluation plans utilized by human service nonprofit organizations. There has been increasing demand for human service organizations to measure performance as a method to account for financial expenditures and to document improved service delivery. Limited research has explored factors related to quality evaluation plan efforts that are conducted in the human service nonprofit sector. Insight into the factors that contribute to evaluation plan quality can make a significant contribution to the existing literature base. This study provided critical information and insight to assist nonprofit sector leaders in understanding the relationship between organizational size, size of funding allocation, type of human service organization and evaluation plan quality. This study also provided insight into the level of evaluation plan quality within the human service nonprofit sector for funding entities that are interested in utilizing evaluation results as a factor when distributing financial resources and in determining the resources, training and technical assistance that can be provided to assist human service organizations in conducting quality evaluations.
Purpose of the Study

The purpose of this correlational study was to examine the relationship between organization size, size of funding allocation, type of human service organization, organization staff position responsible for the evaluation plan and evaluation plan quality in human service nonprofit organizations in Ohio. Evaluation plan quality was assessed by using the Program Accountability Quality Scale (PAQS). This study utilized archival data from 125 human service nonprofit organizations collected by a United Way organization to examine factors related to evaluation plan quality. Nonprofit organizations are defined as entities that have received 501(c)(3) status through the federal government and have a charitable purpose.

Research Questions

1. Does evaluation plan quality significantly relate to organization size for human service nonprofit organizations?
2. Does evaluation plan quality significantly relate to funding allocation size for human service nonprofit organizations?
3. Does evaluation plan quality significantly differ by type of human service nonprofit organization?

Theoretical Framework

The theoretical framework guiding this study was based on evaluation research as conceptualized by Weiss (1998). Evaluation is defined as a formal, systematic assessment of the operations and outcomes of a program for the purpose of improving service offerings. The framework provided by Weiss demonstrated the importance of understanding the purpose and anticipated use of the evaluation. Formative evaluation is concerned with the operations of a program, whereas summative evaluation focuses on program outcomes and benefits for
participants. One of the critical steps in evaluation planning is to determine the value and proposed usage of the evaluation.

Evaluation plans can take many different forms, however effective evaluations have several key components. An effective evaluation includes an evaluability assessment that occurs prior to the evaluation to examine the anticipated program outcomes, the program services and the evaluation. Additionally, when developing the evaluation plan, there are five types of variables that may be measured. These variables include program inputs, unintended program consequences, interim progress measures, program processes and program outcomes. Although each of these variables may not be measured, they should be considered and evaluated as a potential component when designing an evaluation process.

Weiss includes a framework for understanding the role of the researcher or evaluator in the evaluation process. The belief that evaluation quality rigor and complexity can only be accomplished through the use of an external consultant was dispelled. In the proposed model, the researcher can be an active participant as this can lend insight into the program services and increase the practicality and usefulness of the evaluation. Quality evaluations result in information that documents the progress, strength and weaknesses of the program. These evaluations should ultimately result in information that organizational leaders can utilize to make changes or enhancements to existing services. Although an external force often prompts evaluation efforts, the intent of program evaluation will not be fully realized unless the evaluation findings are utilized to inform internal improvement processes.

Significance of the Study

This study resulted in significant contributions to the limited body of literature in the nonprofit field that had explored efforts to examine evaluation plan quality. Findings from this
study were intended to provide insight into the organization factors that relate to evaluation plan quality. While much of the existing literature questions the quality and usefulness of evaluations conceptualized and utilized by nonprofit leaders, few studies explored factors related to evaluation plan quality. This study illuminated organization factors related to evaluation plan quality within the human service nonprofit sector. Findings from this study were intended to provide direction as to where support and resources could be provided to assist organizations in increasing their evaluation plan quality.

Definition of Terms

**Evaluability assessment** – Evaluability assessment occurs before an evaluation system is established. The purpose of an evaluability assessment is to establish a framework to explore the intended outcomes of a program with the actual accomplishments. The evaluability assessment provides an opportunity to assess the perspectives of various stakeholders in regard to the purpose of the evaluation.

**Evaluation** – Evaluation is the process established by human service nonprofit organizations to measure performance related to programmatic outcomes.

**Evaluation plan quality** – Evaluation plan quality is an assessment of the feasibility, usefulness and validity of a system designed to measure program services. The Program Accountability Quality Scale (PAQS) (Poole et al., 2000) was developed as an instrument to assess the quality of evaluation plans in United Way funded human service nonprofit organizations.

**Funders** – Funders represent a range of entities that support nonprofit organizations through financial support, technical assistance and competitive grants. Funders can include organizations such as local United Way organizations, community foundations, private foundations, family foundations, corporate foundations and governmental entities.
**Organization size** – For purposes of this study, organization size was determined using the Internal Revenue Service Form 990. The line item that detailed total annual operating expenses for 2005 was utilized. A scanned copy of the Form 990 was available on the website: www.guidestar.com.

**Program outputs** – Program outputs involve the measurement of the products or services produced by a program. In the nonprofit sector, program outputs often include measures such as counting the number of clients served, number of program participants, or number of meals served.

**Program outcomes** – Program outcomes document impact or change on the well being or health of the participants attributable to participating in or receiving the program’s services. Program outcomes represent the intended result of the program or service. Examples of program outcomes include behavioral changes such as increased knowledge of the effects of smoking, decreased school absenteeism or increased self-sufficiency skills.

**Stakeholders** – Stakeholders include a variety of individuals and entities that have a vested interest in the services and outcomes of an organization. Stakeholders can include funders, individual donors, staff members, volunteers, recipients of service, community officials, and the general public.

**Delimitations**

This study was limited to human service nonprofit organizations that had received funding from a United Way organization located in Ohio. Due to the intensity of understanding and completing evaluation plans, non-United Way human service nonprofit organizations were not considered for participation in this study.
Limitations

A limitation of this study was that the sample was derived from organizations that were funded by the same United Way organization. Regional United Way organizations are affiliated with the national organization, United Way of America, and therefore are required to adhere to certain standards. However, there were differences among United Way organizations in regard to their funding focus, distribution of funds and the amount of training and technical assistance that was offered to human service nonprofit organizations.

Another limitation of this study was that factors, other than those that were explored, may have significantly influenced evaluation plan quality. These factors may have included resources available throughout the nonprofit sector such as opportunities for intensive training and/or technical assistance from other funding sources or the use of external consultants specifically hired to design and implement an evaluation plan.
CHAPTER II. LITERATURE REVIEW

This chapter provides a summary of the literature related to evaluation plan quality within the human service nonprofit sector. Specifically, the following areas were reviewed: understanding the nonprofit sector, the demand for accountability and evaluation, influence of business evaluation on the nonprofit sector, relationship between nonprofit organizations and funders, the use of evaluation systems, development of program based nonprofit evaluation systems, concerns with nonprofit evaluation, organization factors related to evaluation and assessing evaluation plan quality.

Understanding the Nonprofit Sector

Nonprofit organizations are defined as organizations that have received tax-exempt status from the Internal Revenue Service. More than 1.9 million nonprofit organizations have been created to fulfill a range of social, charitable, and religious purposes. The nonprofit sector contributes an estimated 7-11% of the United States gross domestic product (Wagner, 2008) and has experienced steady growth in terms of the amount of services provided, contributions received and number of individuals employed (Light, 2000). Approximately 1.4 million of these nonprofit organizations have 501(c)(3) status ("Facts and figures about charitable organizations", 2007).

501(c)(3) organizations are differentiated from other types of nonprofit, tax-exempt organizations in that they must benefit the public interest through charitable provision of services. These types of nonprofit organizations are referred to by terms such as non-governmental organizations, charitable organizations, civil sector society, or the independent sector as they are differentiated from the government or business sector ("Facts and figures about charitable organizations", 2007; McHargue, 2003).
The Internal Revenue Service further divides charitable organizations into ten categories: arts, culture and humanities; education; environment and animals; health; human services; international and foreign affairs; public and societal benefit; religion-related; mutual membership/benefit and unknown/unclassified (Czerwinski, 2007; "Facts and figures about charitable organizations", 2007). Human service nonprofit organizations include service providers such as homeless shelters, youth development, job training, crime prevention, soup kitchens, recreation and sports. Human service nonprofit organizations represent a crucial aspect of society as they are often responsible for providing a range of services from basic needs such as food, shelter and clothing to more advanced services such as mental health and substance abuse counseling. In the last several decades, federal and state governments have reduced the amount of direct services provided by governmental organizations. Although the government has reduced the amount of direct services provided, they have provided a source of financial support for the nonprofit sector. Due to this reduction in direct services, countless human service nonprofit organizations have filled this void by either expanding services or new organizations have been created to meet the growing need for social services (Czerwinski, 2007; Hammock, 2006).

Demand for Accountability and Evaluation

The widespread use of accountability within the nonprofit sector dates back to the 1960s. Accountability within the nonprofit sector from the 1960s until the 1990s was primarily concerned with the financial accountability of organizations (Plantz et al., 1997; Thayer & Fine, 2001). Accountability was defined as a responsibility to donors, funders and the government to ensure that money was spent in a manner that was consistent with the organization’s mission (Hoefer, 2000). The focus on accountability was initially driven by requirements imposed on the
sector by the federal government as one of the conditions of receiving the benefits afforded nonprofit organizations (Light, 2000). Fiscal accounting of funds was reported through several methods including financial audits, government forms and annual reports produced by the organization (M. Lee, 2004).

Since the early 1990s, the concept of accountability in the nonprofit sector was expanded beyond fiscal reporting to include a programmatic assessment of service provision. While reporting on financial expenditures was sufficient to meet government requirements to maintain tax-exempt status, organizational stakeholders such as funders, donors and the general public pushed the sector to move beyond financial reporting to include an accounting of services (Keating & Frumkin, 2003; M. Lee, 2004). This accounting of program services was often conceptualized as reporting on the “products delivered and people served” (Plantz et al., 1997, p. 2). This typically consisted of programmatic outputs such as reporting the number of clients served, the number of meals distributed or the hours of service provided.

In the 1990’s, the nonprofit sector was once again pushed to further the current practices of accountability and evaluation. The nonprofit sector moved beyond accounting for programmatic outputs to a more substantial analysis and reporting of programmatic outcomes. Program outcomes were defined as a measure of the changes produced by participation in program-related activities (Connolly & York, 2002; Cutt & Murray, 2000; Plantz et al., 1997). These types of outcomes included measures such as increased knowledge of the effects of smoking, decreased school absenteeism and increased self-sufficiency skills. Most human service organizations were routinely measuring and reporting fiscal spending, basic program outputs and demographic information of those served; however, it was uncommon for organizations to measure the impact services had on clients’ well being or the level of client
satisfaction with services (Plantz et al., 1997). Outcome measurement was an effort to document the impact of program services on the participants or recipients of service. Outcomes were defined as “benefits or changes in participants’ knowledge, attitudes, values, skills, behavior, condition or status” (Plantz et al., 1997, p. 3). Documenting program outcomes was intended to provide a more substantial and meaningful measure of the services provided. Conducting programmatic outcome measurement and analysis became a key component to creating an effective and well respected human service organization (Connolly & York, 2002; Cutt & Murray, 2000).

Influence of Business Evaluation on the Nonprofit Sector

Recognizing that evaluation research in the nonprofit sector is in its infancy (Bielefeld, 2006), it is important to acknowledge the influence of evaluation from the business sector. Nonprofit sector evaluation and management research has focused on charitable giving, volunteerism, and comparisons to evaluation within the business sector. Although there is a dearth of research on these topics, a growing body of literature has addressed the use of different types of evaluation systems and their impact on the nonprofit human service nonprofit sector (Bielefeld, 2006; Hoefer, 2000).

While nonprofit evaluation systems may be informed by practices in the for-profit sector, research has demonstrated that it is not feasible to directly apply for-profit evaluation systems to the nonprofit sector (Bielefeld, 2006; Chetkovich & Frumkin, 2003). For-profit evaluation systems are typically not as complex as nonprofit evaluation systems. For-profit systems are often driven by a clearly defined group of stakeholders and agreement on the method utilized to assess performance which typically is done through financial measures (Speckbacher, 2003).
For-profit evaluation models and programmatic decisions are often driven by the profitability of a product or service.

Nonprofit evaluation systems are more complex as they have varying stakeholders (board, employees, clients, funders, public) that may have different perspectives on the value and necessity of services provided. Nonprofit organizations must also take into account the social and mission driven aspects of the human service organization along with the financial measures (Krug & Weinberg, 2004; McHargue, 2003; Wholey, Abramson, & Bellavita, 1986). For example, one of the organization’s programs may not generate sufficient revenue to cover the operating expenses. However, if the program was central to the mission of the organization it therefore would be considered a critical service and may be continued. Although the financial impact of the program may be a consideration when evaluating services, it is not necessarily the driving force.

Consequently, the human service sector has adapted and developed evaluation systems that take into account the perspective of multiple stakeholders, along with sensitivity to the mission of the organization. Although most human service organizations are required by one (or in many cases) multiple funders to engage in a variety of evaluation efforts, there are some organizations that appreciate the value in assessing service delivery beyond the requirements and have therefore voluntarily conducted evaluations. Often times, human service organizations, regardless if the evaluation is voluntary or mandated, find several uses for evaluation results. Information from program evaluation results can be used to assist in securing additional funding or grants, raising funds from individual donors and increasing public confidence in the organization (Cutt & Murray, 2000).
The nonprofit sector has been hesitant to adopt business practices, often citing the inherent differences between organizations formed to fulfill a human service need and those established to generate profits. Another common reason that nonprofit organizations state that they are not able to adopt business practices is due to a lack of resources to implement the management and organizational changes (Beck, Lengnick-Hall, & Lengnick-Hall, 2008). Therefore, nonprofit organizations have often partially adopted business practices with the intent of improving a portion of the organization’s functioning or capacity. While this practice has been commonplace, Beck et al. (2008) completed a study examining the transfer of business practices to nonprofit organizations. This study used a participatory ethnography approach to examine the applicability of business management concepts in a small, nonprofit organization located in the Midwest.

Findings from this study revealed that the transferability of business practices were significantly impacted by two factors: 1) knowledge and expertise of the business practice to be adopted and 2) a thorough understanding of and most importantly respect for the nonprofit environment in which the practice was applied. Based on these findings, the researchers highlighted the importance of these aspects on the successful transferability of business concepts. These factors include that a large number of business practices are adopted (as opposed to implementing one business practice in isolation), the culture and structure of the organization is flexible and adaptable and the organization’s key constituents (employees, board, funders, clients) are invested in the implementation process. While these findings have limitations due to the research approach that was utilized, they appear to correspond with other findings that have called for a sensitivity to the unique aspects of nonprofit organizations when transferring
business practices (Bielefeld, 2006; Chetkovich & Frumkin, 2003; Krug & Weinberg, 2004; McHargue, 2003; Wholey et al., 1986).

**Relationship between Nonprofit Organizations and Funders**

While nonprofit organizations serve a unique role in society, they also receive funding through a unique process. Nonprofit organizations receive financial support through a combination of governmental sources, fees for service and charitable donations. Charitable giving in the United States is comprised of the following: donations from individuals (75%), foundations (13%), bequests (7%) and corporations (5%) (*Giving USA 2008*, 2008). Although foundations do not comprise the largest funding source when compared to governmental sources or even the largest funding source solely among charitable giving, it has been acknowledged that foundations have significant influence on the nonprofit sector (Gronbjerg, Martell, & Paarlberg, 2000). Part of this influence goes beyond the granting of money to include an affiliation with particular funders that gives credibility and community status to an organization (Fairfield & Wing, 2008).

The relationship between nonprofit organizations and funders contains an inherent power differential. One group has financial resources to distribute (funders) and the other group is seeking resources (nonprofit organizations) (Fairfield & Wing, 2008). While some nonprofit researchers have called for funders and nonprofit organizations to develop collaborative relationships (Ostrander & Schervish, 1990), research has documented that this can be difficult to achieve.

A study completed by Fairfield and Wing (2008), explored the nature of the relationships between foundation staff members and nonprofit organizations that had received financial support. Through a combination of focus groups and semi-structured interviews, Fairfield and
Wing explored the transactions that occurred between these two groups. Findings from this study demonstrated that a personal relationship was a key aspect of formulating a relationship that was collaborative and allowed for an honest discourse. Although a collaborative relationship is valued within the sector as it contributes to higher quality programs, evaluation and better use of funding, it is often difficult to achieve.

Recognizing the complex relationship between funders and nonprofit organizations, several research studies have documented concerns among nonprofit leaders regarding how they perceive they have to present information in a particular way to receive funding. Brown and Garg (as cited in Fairfield & Wing, 2008) documented that nonprofit leaders have commented that at times they say “what they think the funder wants to hear in order to avoid conflict and bolster their chances for grant approval” (p. 29). While this lack of transparency has at times been acknowledged by foundation staff, there has not been a significant movement within the foundation sector to improve these relationships.

Contributing to the complexity of the relationship between funders and nonprofit organizations is the reality that funders have established unique funding processes and grant making models. A study completed by Gronbjerg et al. (2000) examined the funding process utilized by various foundations. Results from this study found that foundations included in this study were diverse; however, there were some important findings that impacted the relationship between funders and nonprofit organizations. This study revealed that the diversity among foundations and other sources of funding for nonprofit organizations contributed to a highly fragmented system of supporting human service organizations. Furthermore, funders do not have a common set of values, priorities or perceptions. In fact, funders in this study were unaware of the grant process at fellow foundations. Based on these findings, this study highlights the critical
role of the personal relationship between foundation staff and nonprofit staff in securing grant funding. Instead of instituting a highly rigorous grant proposal review and evaluation process, foundations in this study relied on their personal relationships with nonprofit organization leaders as a determinate in the funding process.

Use of Evaluation Systems

Demand for increased accountability within the nonprofit human service sector resulted in increased usage of formal evaluation systems. Evaluation systems are conceptualized on four different levels: individual, program, organization or community (Cutt & Murray, 2000). Individual evaluation is concerned with the performance of individuals within an organization. Program level evaluation is specific to a sub-section of the organization that offers a particular service or product. Organizational evaluation involves an assessment of the total functioning of the entity. Community based evaluations involve the assessment of a collective community impact on a particular social need or issue.

Within the human service nonprofit sector, evaluation is most typically conceptualized in two ways: as a program evaluation system in which program specific indicators are identified and tracked or a community based evaluation system in which the organization relates its performance to a set of agreed upon community outcomes (Campbell, 2002; Cutt & Murray, 2000; Stoecker, 2007). While nonprofit managers value the importance of striving to impact community outcomes, human service organizations typically utilize a program based evaluation system, as there are a number of complex issues that inhibit the establishment of a community based evaluation system.

However, there are a few communities that have taken on the challenge of developing a community based outcome indicators evaluation (Campbell, 2002). These evaluation systems are
intended to provide stakeholders with comprehensive measurements of various social issues such as high school drop out rates, percentage of uninsured residents, rate of homelessness, and rates of criminal activity. Many communities that have explored this concept have found it difficult to reach consensus on establishing a group of community indicators that could be utilized by stakeholders to assess progress in human service delivery (Campbell, 2002). Furthermore, there are concerns about relating the outcomes of a specific human service program to community indicators in which a number of factors such as other programs, economics, and demographic trends may impact the outcome (Campbell, 2002; Crook, Mullis, Cornille, & Mullis, 2005). Therefore, most human service organizations implement a program based evaluation system.

*Development of Program Based Human Service Evaluation Systems*

Program based evaluation systems can be instituted through several different models that are utilized within the nonprofit sector. Bozzo (2000) identifies three categories of evaluation systems used in the nonprofit sector that strive to respond to the complexities of a human service organization within the evaluation: 1) balanced scorecard, 2) participatory, empowerment and collaborative models and 3) outcome measurement models.

Balanced scorecard most closely resembles evaluation practices within the for-profit sector as business principles are utilized to evaluate service provision. Balanced scorecard evaluations utilize business language and concepts to assess the performance of the organization. The performance evaluation takes place within four organizational domains: financial, customers (clients), internal functioning and innovation/learning (Kaplan & Norton as cited in Caudle, 1997). The evaluation of these four domains provides a systematic assessment of service delivery and the internal functioning of the organization. Although business sector professionals who
volunteer in a leadership capacity for human service organizations have supported this approach, there are concerns about the emphasis on efficiency.

Participatory, empowerment and collaborative models represent an evaluation approach that engages a variety of stakeholders in the evaluation process, which includes designing the evaluation, collecting data, analyzing data and reporting. Participatory models are often considered to be a “bottom-up” approach that allows for a range of input from direct service staff, volunteers, management and board members.

Outcome measurement models, otherwise known as program logic models, were used in the human service nonprofit sector when they were required by United Way of America and local United Way chapters throughout the United States in the mid-1990s (Julian & Kombarakaran, 2006; Newcomer, 1997). According to Bozzo (2000), outcome measurement and logic models demonstrated the relationships between the resources dedicated to a program, the participants, the use of the program resources and then the results achieved by the program. Outcome measurement models typically contained a description of the program resources, inputs, outputs and outcomes (Bozzo, 2000; Connolly & York, 2002; Plantz et al., 1997). Outcome measurement models were considered to be a top-down approach to evaluation since the activities and services that were measured were usually driven by the request of a funding entity or top level management.

While human service organizations have begun to infuse evaluation into daily and strategic planning, it is important to recognize the limitations that currently exist regarding the nonprofit sector’s usage of this information. Since the movement towards evaluation was imposed upon the human service nonprofit sector by a variety of funders, and typically not the organizations themselves, it is understandable that the use of evaluation findings has not
progressed beyond fulfilling funding demands (Cutt & Murray, 2000; Fine, Thayer, & Coghlan, 2000; Light, 2000).

*Concerns with Nonprofit Evaluation*

While there has been documented increase in the use of evaluation systems, the implementation of these systems has been questioned by a variety of stakeholders including funding entities, nonprofit sector leaders and at times the public (M. Lee, 2004; Light, 2000; Murray, 2001; Newcomer, 1997). Considering that evaluation systems utilized by human service organizations are adapted by each organization in an effort to evaluate its unique services, it is difficult to assess, in a systematic manner, the capacity of the evaluation systems to accurately measure service results (Poole et al., 2001). Furthermore, human service organizations have been established to address complex societal issues and evaluating the success of human service organizations can be problematic (Gronbjerg et al., 2000). However, research efforts within the human service nonprofit sector have explored a variety of evaluation components including the use of evaluability assessment, the role of stakeholders, and the usefulness of existing data collection efforts (Campbell, 2002; Poole et al., 2000; Stoecker, 2007; Weiss, 1998).

*The Use of Evaluability Assessment.*

While the human service nonprofit sector has adapted to the demands of conducting evaluations, there are concerns regarding the current capacity of human service organizations to implement evaluation systems. These concerns relate to the organization’s ability to focus on the quality of the evaluation, available resources and the usability of the data. An important component of establishing a successful evaluation system is that an evaluability assessment occur, which “increases the likelihood that a proposed measurement system will deliver what it promises.” (Poole et al., 2000, p. 16). An evaluability assessment is a critical component that
ideally takes place before the evaluation system is established as it provides a framework to explore the intent of a program with the actual accomplishments. Instituting an evaluability assessment provides an opportunity for a range of stakeholders to participate in the design of the evaluation prior to evaluation activities taking place. Involving stakeholders throughout the evaluation process allows for more effective and higher quality evaluations. Using an evaluability assessment forces nonprofit leaders to conceptualize the ways in which an evaluation can be developed to improve service delivery (Weiss, 1998; Wholey & Newcomer, 1997).

The Role of Stakeholders.

Stakeholders have a unique role in the function of nonprofit organizations as they have varied roles, responsibilities and influence on the organization. In human service organizations, stakeholders may include funders, individual donors, staff members, volunteers, recipients of service, community officials, and the general public. Stakeholders critically influenced the role of accountability in the nonprofit sector through their demands that organizations complete evaluations. There have also been growing expectations among the public and general donors that organizations demonstrate the effectiveness of their services. Several researchers have explored the influence and involvement of stakeholders in nonprofit evaluation.

Research completed by Thayer and Fine (2001) examined the role of stakeholders in the evaluation systems and practices of nonprofit organizations. The purpose of the study was to explore the types of evaluations completed by nonprofit organizations and to gain deeper understanding into the perceived usefulness, credibility and satisfaction of the evaluation related to the level of stakeholder involvement. This study consisted of 140 nonprofit organizations that provided services and had completed a program evaluation in the last three years. Participants
completed a quantitative survey and of those participants, forty participated in a follow-up telephone interview. Results from this study found that 50% of the respondents engaged in evaluation efforts because of a funding requirement and approximately 20% conducted evaluations because they to evaluate whether or not their programs had an impact on the clients they served.

The involvement of stakeholders in the various aspects of the evaluation process had an impact on the perception of the evaluation. Thayer and Fine identified four categories of stakeholders: nonprofit staff members, board members, staff members at funding entities and participants or clients that benefited from the organization’s services. Stakeholders contributed in unique ways to the evaluation process. Nonprofit staff members were most involved in designing the evaluation, collecting and analyzing data and writing reports. Board members and funder staff members focused on designing the evaluation and interpreting the data. For organizations that participated in this study, participants or clients typically did not contribute to the evaluation process.

Analysis of the data found that nonprofit organizations with high levels of stakeholder participation considered their evaluation to be more useful, credible and satisfying when compared to organizations with lower levels of stakeholder participation. Evaluations that are useful, credible and satisfying “have a solid and focused design, document programmatic success and offer recommendations for program improvement” (Thayer & Fine, 2001, p. 108). This study demonstrated the important role that stakeholders can have in designing, collecting and analyzing organizational evaluation efforts that are of benefit to nonprofit organizations and the clientele they serve.
A study completed by Cutt and Murray (2000) specifically examined the relationship between funders and human service organizations that had been funding recipients. This research involved a qualitative, grounded theory approach to explore the evaluation process. Three major funders and four human service organizations that received funding from these three entities were included in the study. Research findings demonstrated that not only did the funders expect different evaluation information; they also expected evaluations to be conducted at different system levels. Collectively, the funders requested evaluation information related to program, organization and community level results. Although the qualitative interviews highlighted the intentions of funders to utilize outcome measurements in their funding decisions, further examination found that the funders mainly utilized resource and output measures.

**Usefulness of Evaluation and Data Collection Efforts.**

Recognizing the growing role of data collection and management for nonprofit organizations, Stoecker (2007) conducted a survey that assessed the research and data needs of nonprofits. Small to medium sized nonprofit organizations were selected to participate in the study. Eighty nonprofit human service organizations in a metropolitan area completed the survey. Results from this study demonstrated that organizations collected large amounts of data and dedicated, on average, 56 hours per week to data collection and management. The data that was most frequently collected was client demographic information such as age, sex, race/ethnicity, address and family characteristics. While resources were dedicated to data management, Stoecker found that the data was typically not used by organizations. Many organizations reported that they did not utilize the data, beyond fulfilling funding or legal requirements. Barriers that were identified to utilization of the data was that the type of data
collected was not perceived as useful and the organizational staff members did not have the skills or research capacity to more fully integrate the data into service delivery.

Additional research completed by Broom and Jackson (1997) found that increasing the research and evaluation capacity of human service organizations was critical to increasing the quality and usability of the evaluation, data management, and ultimately the provision of services. Research findings demonstrated that evaluation training must be unique to the skills and needs of the organization’s staff members. Training topics and content should have been varied based on the organizational position of the participant. That is, the training needs for an executive director of an organization typically differ from program management or direct service staff members. Additionally, research and evaluation training designed for the nonprofit practitioner was more effective when conducted by those who had first hand experience conducting evaluation in the nonprofit sector or were knowledgeable and sensitive to the unique challenges faced by human service organizations.

Fine, Thayer and Coghlan (2000) explored the use of program evaluation within the nonprofit sector. Through a mixed methodology research design, participants included 140 organizations that completed a survey. Of those participants, 40 completed a telephone interview and an in-depth case study was completed with four organizations. The research findings demonstrated that the evaluation efforts of nonprofit organizations were considered useful when several of the following factors were present: recommendations for program improvement were included as a component of the evaluation, the evaluation results were utilized as a planning tool, program success was documented, the design of the evaluation was specific and focused, there was a high level of stakeholder involvement and the evaluation was conducted by both internal program staff and external consultants.
These factors identified by Fine et al. (2000) contribute to the level of knowledge usability (Hammock, 2006). In this context, knowledge usability is considered to be the use of evaluation findings to either inform services that are currently being provided or to inform decisions regarding future service provision. The use of the evaluation findings, beyond fulfilling a reporting or documentation requirement for a funding entity, can have a significant impact on the perceived value and quality of the evaluation (Fine et al., 2000; Salipante & Aram, 2003).

Organization Factors Related to Evaluation

The movement toward evaluation within the nonprofit sector has recently prompted the exploration of the relationship between organization factors and evaluation initiatives. Although there have been limited studies exploring organization factors, these studies have provided insight and a framework for further exploration.

Organization Size and Evaluation.

Organization size relates to a variety of measures used to compare the financial, programmatic and capacity aspects of an organization. Organization size variables include measures such as asset size, number and type of service programs, annual income and expenses, and number of employees (Ebaugh, Saltzman Chafetz, & Pipes, 2005; B. R. Lee, McMillen, Knudsen, & Woods, 2007).

A study completed by Robertson (2007) explored relationships between school district size and cost efficiency and performance. This study was completed in response to the movement to consolidate school districts to create economies of scale. Consolidation was intended to create three results: tax payer savings, reduction in duplicative services and improvement in education quality. This study explored each of these anticipated results through a cross-sectional analysis of data from large school districts located throughout the United States.
Results revealed that for the large school districts included in the study that district size did not decrease the amount of funding spent per student, district size did not positively impact education quality and district size did not impact efficiency. Therefore, establishing a large district (organization) did not have a positive impact on the quality of educational services. Robertson explains that the consolidation of school districts based on geographic proximity resulted in diseconomies of scale as the large organizations struggled to respond to local needs.

Further research conducted by B. R. Lee, et al. (2007) explored organizational efforts to increase quality services and barriers to implementing quality services. Organization factors such as organization size, type of organization, accreditation or licensing status and agency’s affiliation were examined. Utilizing an instrument that was developed to assess activities and barriers related to achieving quality services, questionnaires were mailed to field practicum instructors affiliated with a master of social work program in the Midwest that held positions at social service organizations. Results from this study revealed that organizations implemented quality programmatic efforts and most organizations utilized a range of methods. Additionally, accreditation and licensing status were related to quality program activities. Barriers to increasing the quality of services included limited financial resources.

Of particular interest are the findings related to organization size and type of organization. Organization size was determined by the number of employees at the organization (less than 50 employees, 51-100 employees, 101-500 employees, 501-1000 employees and over 1000 employees). Findings from this study revealed that two aspects of quality related activities (quality personnel positions and quality oriented processes) were related to organization size. Quality personnel positions were defined as hiring specialized staff to promote and maintain quality. Quality oriented processes were defined as organizational level policies and procedures
designed to have an impact on offering quality services. Organizations with 101-500 employees and organizations with over 1000 employees were significantly more likely to have quality personnel and quality oriented processes when compared to organizations that had less than 50 employees.

In this study, the organization size variable related to a portion of the quality programmatic outcome measures with a majority of the larger organizations faring better. However, B. R. Lee et al. (2007) recommended further exploration and analysis of the factors related to program quality due to the limitations of the methodology and sample in their study.

*The Role of Funding Level and Evaluation.*

Securing adequate and ongoing financial support has been acknowledged as one of the most difficult and stressful aspects to the work of nonprofit organizations (Brown & Troutt, 2007; Hughes & Luksetich, 2004; B. R. Lee et al., 2007). Although there has been much rhetoric about the importance of evaluation quality and increased accountability in the nonprofit sector, studies have documented the pivotal role that funding level and type of funding in organizations has in engaging in quality evaluation efforts.

Research completed by Brown and Troutt (2007) examined the various characteristics of funding that related to accountability, quality of relationship between the funder and the nonprofit organizations and the financial stability of the organization. The funding characteristics that related to duration, likelihood of funding renewal and restrictions had a significant impact. Duration of funding, which ranged from months to years, most strongly impacted the organization’s financial stability. Likelihood of renewal was the organization’s perception that funding would be renewed during the next cycle or funding opportunity. Likelihood of renewal impacted accountability measures and financial stability. Restrictions on the use of funds
impacted organizational flexibility and the quality of the funding relationship based on the severity of restrictions placed on the funds. Brown and Troutt (2007) also emphasized the negative impact of funding that was characterized by a short duration, was limited in size and was highly restrictive. This type of funding often resulted in lower quality evaluations and program outcomes, due to a lack of sufficient resources that must be dedicated to the evaluation process in order to establish a quality program and evaluation process. In situations such as this, resources are typically spent on program implementation and service provision with little attention given to evaluation. These conclusions were supported by Lee et al. who stated that “resource limitations likely constrain quality efforts. Providing more services and better care with fewer resources characterizes the current context of social work practice” (2007 p. 68).

Research conducted by Carrilio, Packard and Clapp (2003) explored issues related to the usage of a performance management evaluation system. The authors detailed the importance of sufficient funding in utilization of a performance management evaluation system. Barriers to implementing a quality evaluation system are typically due to a lack of knowledge and/or a lack of sufficient funding. The study conducted by Carrilio et al. (2003) examined the implementation of a statewide performance management system designed to evaluate services provided to families. Findings from the study revealed that funding and technical support and training were considerations in utilization of the evaluation system, but there were other factors that had an impact. Staff members’ attitudes and beliefs regarding data and research and the complexity of the evaluation system appeared to have a significant impact. For example, organizations in which the staff members exhibited negatives views regarding research and data were not likely to utilize the evaluation system. Therefore, in order to effectively institute quality evaluations, Carrilio et al. recommend support from external factors such as funding and training along with
support from internal factors such as stakeholder investment and involvement in the evaluation process.

Type of Organization and Evaluation.

Several researchers have recommended further exploration of the relationship between type of human service organization and nonprofit evaluation efforts (Broom & Jackson, 1997; B. R. Lee et al., 2007; Poole et al., 2001). Although it is acknowledged that human service organizations provide a wide range of services to society, there are distinct differences between the various types. While United Way organizations across the country tend to categorize human service organizations in a similar matter, it is important to point out that other funders have utilized different methods of categorizing organizations.

The research study completed by Poole et al. (2001) explored the relationship between agency culture and evaluation plan quality. This study was completed with United Way funded agencies and utilized information gathered from telephone interviews and archival data. Telephone interviews were utilized to assess a variety of factors, including agency culture. Archival data included an assessment of the outcome measurement plans submitted to United Way. The Program Accountability Quality Scale was used to assess the outcome measurement plans. Findings from the study revealed that agency culture did not affect evaluation plan quality. Although agency culture, defined as the organizational beliefs and values, was not significant, the researchers recommended another aspect of agency culture, such as the type of human service organization, be considered as a variable in future research.

Further exploration of the relationship between type of organization and evaluation quality was related to Broom and Jackson’s (1997) study that examined organizational capacity and usefulness of the evaluation system. Broom and Jackson reported several findings related to
the interaction between organizational capacity and evaluation quality. These findings focused on staff position specific recommendations such as offering evaluation training tailored to the skills and needs of staff members. While these findings and recommendations were intended to increase evaluation capacity and quality, the study did not account for type of human service organization.

A study that explored the relationship between type of human service organization and quality services was completed by B. R. Lee et al. (2007). For purposes of this study, type of organization was divided into four categories: mental health organizations, health organizations, child welfare organizations and other (organizations that did not fit the first three categories). Findings from this study revealed significant differences between the health organizations and the organizations that were identified as “other” regarding activities that were implemented to achieve quality programs and outcomes. Health organizations had more quality personnel positions and more quality processes than organizations in the other category. While this study supports the belief that quality processes and evaluation might differ by type of organization, further research was recommended to explore and test these findings.

Assessing Evaluation Plan Quality

The concentrated movement towards accountability in the nonprofit sector resulted in numerous evaluation approaches (Bozzo, 2000; Julian & Kombarakaran, 2006; Murray, 2001; Thayer & Fine, 2001). Most nonprofit organizations dedicated resources to evaluation efforts, however little research explored evaluation plan quality. Understanding evaluation plan quality is of critical importance, as it provides the foundation for the entire evaluation process. A significant aspect of the motivation for stakeholders to be involved in the accountability movement within the human service nonprofit sector was to require that organizations evaluate
their services. The evaluation was intended to ensure that organizations that received funding provided substantial information that demonstrated the ability of the organizations to have an impact on the lives of the participants. Funding stakeholders intended to utilize the evaluation information to determine future funding decisions (Wholey & Newcomer, 1997). If a system such as this were implemented, nonprofit organizations that demonstrated success or impact on outcome measurements would be more likely to continue receiving funding whereas organizations that did not demonstrate success would be penalized financially. Persistent concerns stated by nonprofit organization leaders and funders regarding the quality of evaluation systems has prohibited nonprofit leaders from using outcome evaluation results to determine funding decisions (Julian & Kombarakaran, 2006; Plantz et al., 1997; Poole et al., 2000).

Designed to address concerns regarding the quality of evaluation plans, Poole et al. (2000) developed the Program Accountability Quality Scale (PAQS) as a method for assessing the quality of an evaluation plan system. The PAQS provides a method to complete an evaluability assessment, intended to provide feedback and critical information that will result in a quality evaluation plan system. The PAQS contains seven subscales that relate to the following nonprofit program service areas: resources, activities, outputs, outcomes, goals, indicators, and evaluation. The resources subscale is used to assess the program’s essential components through measuring the resources dedicated to or consumed by a program. The activities subscale assesses the activities associated with the program that are intended to fulfill the organization’s mission. The outputs subscale is related to the direct products of the program and aspects of the program such as how many people are being served and over what period of time. The outcomes subscale measures the accomplishments and benefits of the program. The goals subscale assesses the relation of the program to the larger organization mission and impact of the program. The
The evaluation subscale assesses the reliability, efficiency and realistic expectations of the evaluation system. Evaluation of each of these subscales provides a mechanism to compare the evaluation plan systems of various nonprofits. Research completed using the PAQS provided insight into the quality of the evaluation plan systems, the need for technical assistance and feedback for nonprofit leaders (Julian & Kombarakaran, 2006; Poole et al., 2001; Poole et al., 2000).

Further research efforts by Poole et al. (2001) explored organizational factors that contributed to quality evaluation plan systems in nonprofit organizations. The purpose of this study was to provide nonprofit leaders and funders with research driven information to assist in efforts to conduct quality evaluations. Outcome evaluation systems in United Way funded organizations in a metropolitan area were evaluated to determine the theoretical influence of five predictors of quality: agency culture, management support, technical capabilities, stakeholder involvement and funder approach. In this study, agency culture was defined as the “beliefs, norms, and values that influence the behavior of people who work in nonprofit agencies” (Reisman as cited in Poole et al., 2001, p. 406). Management support was identified as the commitment by top management to support outcome evaluation by creating a successful work environment and establishing clear expectations. Technical capabilities were defined as having appropriate technological capabilities and having staff members knowledgeable about technology. Stakeholder involvement was the degree of inclusiveness and involvement of various staff personnel and board members. Funder approach referred to the range of relationships that funders had with nonprofit organizations from basic granting of funds to a more involved relationship that involved training, technical assistance and consultation.
The United Way funded organizations that participated completed a Likert-scale questionnaire that included statements related to the five predictors of quality. The data was analyzed to examine relationships among the various predictors and evaluation plan quality, which was scored using the PAQS. Findings demonstrated that stakeholder involvement had the strongest direct relationship to evaluation plan quality. Management support also had a strong direct relationship to evaluation plan quality. Management support and technical capabilities had an indirect influence on evaluation plan quality. Findings from this study demonstrated the influential role that stakeholders and management staff had on the role of evaluation within the organization. Stakeholder involvement had the strongest direct relationship, which demonstrated the potential impact that involving a range of stakeholders that have a vested interest in the organization can have in the development, implementation and assessment of the evaluation plan. Additionally, evaluation plan quality was related to the perceived use of technology and staff support. Therefore, organization leaders must “communicate a clear message of the overall importance of outcome evaluation plan to the agency, allocate resources to support planning and implementation processes and build evaluation into performance appraisal systems” (Poole et al., 2001, p. 418).

Results from this study found that agency culture did not affect evaluation plan quality. Agency culture was identified as organizational beliefs and values regarding evaluation. This finding was different than previous research findings that demonstrated that the agency beliefs and values impacted evaluation plan quality. The researchers note that data collected by this study was limited in that there were a small number of participants, only one agency representative completed the questionnaire and each of the five predictors were based on the perception of the factors rather than an independent assessment. Further research is needed to
clarify the relationship between agency culture and evaluation plan quality, the role of funding in evaluation plan quality and the relationship between the complexity of program outcome measures and evaluation plan quality.

The PAQS was used as an evaluation framework for research completed by Julian and Kombarakaran (2006). This research study explored, through a case study approach, the process of implementing an outcome measurement evaluation in a United Way organization. The researchers worked collaboratively with United Way officials to design an evaluation plan quality measure that could be used to assess the outcome measurement plans submitted by organizations the United Way was considering for funding. Utilizing the PAQS criteria, a four-category assessment measure was developed. The four-category assessment measure included the following categories: inadequate – outcome measures were inconsistent with program intent, concerns – weak or insufficient outcome measures, adequate – outcomes were identified and appropriate information was collected, and exemplary – outcomes were well established and information collected was used to support program improvements.

Research findings demonstrated that over a four-year period, outcome measurement assessments improved since more programs were in the adequate and exemplary categories and fewer programs were considered to be in the inadequate category. These research findings demonstrated that organizational capacity in regards to evaluation was improved as the United Way implemented outcomes based measurement. The four-category evaluation plan assessment was considered credible due to several factors. The assessment was based on the PAQS, an evaluation plan instrument that had been empirically tested. A review of United Way funding principles and a review of published literature was completed. Also, both agency and United
Way representatives that were involved in the assessment process received appropriate training and support.

Summary

In summary, the human service nonprofit sector has responded to demands to engage in evaluation efforts and document impact or results that are derived from participation in programs and services. Since the mid-1990s, the nonprofit sector has adjusted to these demands and most organizations have adapted an evaluation plan system. Even though evaluation is more commonplace in the nonprofit sector, concerns regarding evaluation plan quality still remain. However, there has been limited research that has examined organizational factors that relate to evaluation plan quality. In the literature, various factors including the three examined in this study (organization size, size of funding allocation and type of human service organization) have been identified for future research to expand the understanding of the interactions and relationships between organization factors, quality programs, evaluation plans and accountability.
CHAPTER III. METHODOLOGY

This study examined the relationship between organization factors and evaluation plan quality in human service nonprofit organizations. This chapter describes the research design, sample, instruments, procedures for data collection, and data analysis procedures.

Research Design

The design of this quantitative study was correlational. This study examined the relationship between organizational size, size of allocation, type of human service organization, and evaluation plan quality for human service organizations funded by a United Way organization in Ohio. A correlational research design was appropriate for this study because it examined the relationships between variables (Fraenkel & Wallen, 2006).

One of the independent variables in this study was organizational size. The size of the human service organization was determined by using the Internal Revenue Service Form 990. The size of the organization was determined according to the annual operating expenses reported for 2005. The relationship between organizational size and evaluation plan quality was of interest due to the vast differences in operating size and therefore the resources, programs and services that exist among human service organizations. Within the nonprofit sector, human service organizations include large organizations such as mental health service providers and hospitals with budgets of over $1,000,000 to small organizations such as church-based food pantries and after school programs operated by volunteers and donations that have operating budgets less than $25,000.

The second independent variable that was examined was funding allocation size. United Way provided funding allocations to human service organizations that ranged from $5,400 to $600,000. Funding allocations were determined through an allocations committee that reviewed
proposals and made funding decisions. Funding allocation size was determined based on the amount requested by the organization, the size and type of program and ultimately the amount of available United Way funds. Therefore, it was important to examine the relationship between the size of funding allocation and evaluation plan quality.

The third independent variable that was examined was the type of human service organization. It is the mission of United Way of America and local United Way affiliates to provide resources to human service organizations that impact the lives of individuals and families. Organizations self-identify the category of human service that most closely relates to their mission and programming when submitting their allocation funding request based on the target population and intended outcomes for the proposed program. This categorical variable was distributed into categories of human service organizations: programs for young children, programs that served children and teens, and programs that empowered vulnerable families. The type of human service organization was explored based on previous research completed by Poole, Nelson, Carnahan, Chepenik and Tubiak (2000), Broom and Jackson (1997) and Lee et al. (2007) that identified a need to explore the relationship between type of organization and evaluation plan quality.

It was anticipated that this study would include a fourth independent variable, organization staff position responsible for the evaluation. However, the researcher learned that this information was not tracked by the United Way organization and therefore, it was not possible to include this variable.

The quantitative dependent variable was evaluation plan quality, which was measured by the Program Accountability Quality Scale (PAQS). The PAQS was developed specifically for use in the nonprofit sector to quantitatively examine the quality of evaluation plans being
submitted by United Way funded nonprofit organizations. The PAQS utilized data reported in an outcome plan submitted by a human service organization to determine a score that represents the level of evaluation plan quality. The PAQS consisted of the following seven subscales: resources, activities, outputs, outcomes, goals, indicators and evaluation. There were five subscales (resources, activities, outputs, outcomes, and goals) that were combined to create the logic model score. To create the PAQS score, scores from the seven subscales were combined to create an overall measure of evaluation plan quality.

*Sample*

The sample for this study was comprised of human service nonprofit organizations that received funding from a United Way organization in Ohio in 2006. This sample was accessed by using archival data collected by United Way during their funding allocation review process from 2006-2008. The human service organizations submitted an outcome measurement plan to the United Way as a condition of their funding allocation. There were 125 organizations that received a funding allocation from United Way and these organizations represented a range of human service nonprofit organizations. For purposes of this study, archival data from each of the 125 organizations comprised the sample.

In 2006, the United Way organization offered workshops regarding outcome measurement. Also, when requested by human service nonprofit organizations, United Way officials provided individualized guidance and technical assistance as organizations completed their outcome measurement plan. Participation in these opportunities (workshops and technical assistance) was not mandatory.

The United Way organization that participated in this study served a three county area that covered urban, rural and suburban areas. The United Way had an annual operating budget of
$14,000,000 and the mission of the organization was to have a positive impact on the community by providing resources to assist individuals, families and children in building a better life for themselves.

**Instruments**

The independent variables that included organization size (based on annual operating expenses), size of funding allocation and type of human service organization were gathered by reviewing data from two sources. The organization size was collected through examination of the Internal Revenue Service Form 990 that was catalogued on the website: www.guidestar.org. The size of funding allocation and type of human service organization was information that was collected by and stored at United Way. The size of funding allocation and type of human service organization was provided to the researcher in a spreadsheet that detailed this information for each organization that received funding in 2006. United Way allocated funding to human service organizations for a three-year period. Therefore, the opportunity to submit an allocation funding form occurred approximately once every three years. Human service organizations that received funding in 2006 were considered for participation in this study.

The dependent variable, evaluation plan quality, was assessed by examining outcome information submitted to United Way. Outcome information was submitted to United Way six times during the three-year funding period. At the beginning of every funding year (2006, 2007, 2008) organizations were required to submit an outcome plan that detailed the projected resources, activities, outputs, outcomes and goals. At the end of the funding year (2007, 2008, 2009) organizations were required to submit outcomes results that detailed the actual resources activities, outputs, outcomes and goals achieved by the program in relation to the targeted or anticipated program results. Organizations were required to submit an outcome plan and
outcome results in a timely manner during the established time period. Failure to comply with these requirements resulted in United Way withholding funding until the appropriate documents were submitted.

Evaluation plan quality was measured by the Program Accountability Quality Scale (PAQS). This instrument was specifically developed in response to a need identified by nonprofit sector leaders to develop reasonable and useful methods to measure the quality of evaluation plans of nonprofit organizations. Outcome measurement systems were considered to be a prime source for exploration of evaluation plan quality for two main reasons: 1) outcome measurement was widely used by United Way organizations throughout the United States which equated to thousands of programs have adopted this evaluation method and 2) while the content information varied from program to program, the components included in the outcome measurement were relatively the same. Therefore, the structure of this evaluation lent itself to the establishment of a systematic assessment of evaluation plan quality.

Through a collaborative effort, researchers and United Way leaders developed the Program Accountability Quality Scale (PAQS) to explore the quality of evaluation plans submitted by United Way funded organizations (Poole et al., 2000). The PAQS was developed over a four-year period in which training and workshops were offered to United Way funded organizations to assist in the creation of an evaluation system. During this period, the PAQS instrument was developed based on performance measurement theories. Seven assessment areas were identified through analysis of the performance measurement theories. Initially, the PAQS was piloted with 191 programs that were funded at 78 nonprofit organizations. The instrument consisted of seven subscales: resources, activities, outputs, outcomes, goals, indicators, and evaluation plan and contained a total of 34 items. Two evaluators conducted the assessments.
Scoring the outcome measurement plans using the PAQS took approximately 25 minutes. Interrater reliability was determined by selecting a random sample of cases that were rated by three researchers. The interrater reliability ranged from .81 to .92.

At the completion of the pilot study completed by Poole et al. (2000), thirteen items that were not strongly related and two of the subscales were eliminated. Analysis of the overall consistency of the revised scale demonstrated good internal consistency. Reliability for the subscales was also good.

Poole et al. (2000) created a scoring method for the PAQS by analyzing the archival outcome plan data. Each of the 21 items was given a score ranging from 1-strongly agree/very good performance, 2-agree/good performance, 3-disagree/fair performance, 4-strongly disagree/poor performance. Based on this scoring method, lower scores equated to higher quality evaluation plans and higher scores equated to lower quality evaluation plans. Scores on each of the items were added together to determine the subscale score. Subscale scores were then combined to determine the PAQS score.

Procedures for Data Collection

Approval was secured from the Human Subjects Review Board at Bowling Green State University prior to initiating the research study. Informational meetings took place with the Chief Operating Officer and Data Management Specialist at the regional United Way to inform them of the nature of this research project and answer questions. Information and archival data gathered through this research project was coded to ensure confidentiality.

Organization demographic information that included the size of allocation funding and type of human service organization was gathered from spreadsheets generated by the United Way office. The researcher collected this information from electronic copies of the spreadsheets,
which detailed this information for 2006. The other organization factor, organization size, was
gathered from the website: www.guidestar.org. The annual operating expenses total figure was
utilized as the measure of organization size.

Outcome measurement plans were scored using the PAQS. The researcher utilized paper
copies of the outcome measurement plans collected by the United Way to assess the quality of
the plans. All of the analysis of the outcome measurement plans and scoring took place at the
local United Way offices. The outcome measurement plans were scored according to the method
developed by the research team that created the PAQS instrument. In order to determine
reliability, a test-retest method was utilized with a percentage of the outcome measurement
plans. The test-retest method consisted of administering the same instrument to the same group
after a specified period of time (Fraenkel & Wallen, 2006). This method of reliability was
conducted with eight randomly selected organizations. The researchers scored the outcome
measurement plan for the eight organizations, then after a two-week interval the researchers
scored the randomly selected outcome measurement plans again. The test-retest analysis
demonstrated a reliability coefficient of .91.

Data Analysis Procedures

Data was analyzed using the Statistical Package for the Social Sciences. Descriptive
statistics including means and standard deviations were conducted and reported for each variable
and for group comparisons. Data generated from the test-retest process was analyzed by using a
reliability coefficient.

For the first research question (Does evaluation plan quality significantly relate to
organization size for human service nonprofit organizations?) inferential statistics were used.
The Pearson product-moment coefficient of correlation was used to examine the relationship
between the size of the organization and evaluation plan quality. The Pearson correlation was appropriate for this data analysis because both the independent and dependent variables were quantitative measures (Fraenkel & Wallen, 2006). The researcher also examined organizational size group differences in evaluation plan quality using analysis of variance. Three categories (small, medium and large) for organization size were created. These categories were identified by examining a frequency histogram. The categories were created in an effort to generate reasonable groups of a similar sample size.

Inferential statistics were also used to analyze the data collected for the second research question (Does evaluation plan quality significantly relate to funding allocation size for human service nonprofit organizations?). The Pearson product-moment coefficient was used to examine the quantitative independent variable, size of funding allocation, and the quantitative dependent variable, evaluation plan quality. Analysis of variance was utilized to explore funding allocation size group differences. These categories were identified by examining a frequency histogram with the intent to create groups of similar sample size. Post hoc analysis was completed using the Bonferroni technique to examine differences between the groups.

The third research question (Does evaluation plan quality significantly differ by type of human service nonprofit organizations?) was analyzed through an analysis of variance. Analysis of variance was appropriate for this research question as examined if there were significant differences between the means of two or more groups (Fraenkel & Wallen, 2006). For significant results, post hoc analysis was completed using the Bonferroni technique to examine differences between the various groups.
CHAPTER IV. RESULTS

Introduction

The purpose of this quantitative, correlational study was to examine the relationship between organizational size, size of funding allocation, type of human service organization and evaluation plan quality in human service nonprofit organizations in Ohio. This study utilized archival data from the evaluation plans of 125 human service nonprofit organizations collected by a United Way organization to examine factors that relate to evaluation plan quality. The Program Accountability Quality Scale (PAQS) instrument was utilized to assess evaluation plan quality.

The results were presented for each research question. First, descriptive statistics were presented. Second, a correlation that examined the relationship between organization size and evaluation plan quality was presented along with an analysis of variance that examined organization size groups and evaluation plan quality. Next, a correlation that examined the relationship between funding allocation and evaluation plan quality was presented. An analysis of variance examined the relationship between funding allocation groups and evaluation plan quality. Finally, an analysis of variance was presented that examined the relationship between type of human service organization and evaluation plan quality.

Sample Characteristics

This study utilized archival data from 125 human service nonprofit organizations collected by a United Way organization. There were a total of 125 programs that received a funding allocation from United Way in 2006, 109 completed and submitted an evaluation plan. Sixteen programs did not complete an evaluation plan; 11 of those were programs that United Way contracted for services and therefore were not required to submit an evaluation plan; 3 were
programs that were combined into other similar programs; one organization was no longer in business and one program was missing an evaluation plan. There were 15 programs that provided services for young children ages birth to five, 41 programs that served children and teens and 53 programs that were intended to support and empower families.

In Table 1, descriptive statistics representing the level of organization size, funding allocation, and the PAQS subscales (resources, activities, outputs, outcomes, goals, indicators and evaluation) were represented. The mean organization size mean was $3,714,019.47 with a standard deviation of $5,108,144. The minimum organization size was $112,999 and the maximum was $25,932,720. The mean allocation size was $72,135.41 with a standard deviation of $77,363.22. The minimum allocation was $5,400 and the maximum was $600,000.

Items on the PAQS were scored using the following scale: 1 = strongly agree/very good performance, 2 = agree/good performance, 3 = disagree/fair performance, 4 = strongly disagree/poor performance (see Appendix A). It is important to note that for this instrument, lower scores equated to higher evaluation plan quality and likewise, higher scores equated to lower evaluation plan quality.

The resources subscale contained three questions related to the resources dedicated to or consumed by the program. The resources subscale had a possible score between 3 (strongly agree/very good performance) to 12 (strongly disagree/poor performance). The resources subscale mean score was 6.78 with a standard deviation of 1.49.

The activities subscale contained two questions, which assessed the activities undertaken by the program to fulfill its mission. The possible activities subscale score was between 2 (strongly agree/very good performance) to 8 (strongly disagree/poor performance). The activities subscale mean score was 3.58 with a standard deviation of 1.40.
The outputs subscale contained three questions related to the direct products of the program activities. The outputs subscale had a possible score between 3 (strongly agree/very good performance) and 12 (strongly disagree/poor performance). The outputs subscale mean score was 6.75 with a standard deviation of 1.38.

The outcomes subscale contained three questions that examined the benefits for program participants and the measurable accomplishments of the program. The outcomes subscale had a possible subscale score between 3 (strongly agree/very good performance) and 12 (strongly disagree/poor performance). The outcomes subscale mean score was 6.05 with a standard deviation of 1.90.

The goals subscale contained two questions related to the mission of the organization and the larger impact of the program. The goals subscale had a possible score between 2 (strongly agree/very good performance) and 8 (strongly disagree/poor performance). The goals subscale mean score was 6.06 with a standard deviation of .91.

The logic model variable was the sum of the resources, activities, outputs, outcomes and goals subscales. The possible logic model score was between 13 (strongly agree/very good performance) to 52 (strongly disagree/poor performance). The logic model mean score was 29.22 with a standard deviation of 4.34.

Additional PAQS subscales included the indicators subscale, which consisted of four questions related to the outcome measurement process. The indicators subscale had a possible total score between 4 (strongly agree/very good performance) to 16 (strongly disagree/poor performance). The indicators subscale mean score was 8.98 with a standard deviation of 3.09.

The evaluation subscale was comprised of four questions related to data collection and the evaluation process for a possible score between 4 (strongly agree/very good performance) to
16 (strongly disagree/poor performance). The evaluation subscale mean score was 7.58 with a standard deviation of 2.13. The PAQS score was calculated by adding the logic model score, along with the indicators and evaluation subscale scores.

The possible PAQS range was between 21 (strongly agree/very good performance) to 74 (strongly disagree/poor performance). The PAQS mean was 46.06 with a standard deviation of 7.70.

Table 1

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Minimum</th>
<th>Maximum</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organization size</td>
<td>92</td>
<td>112999</td>
<td>25932720</td>
<td>3714019.47</td>
<td>5108144.64</td>
</tr>
<tr>
<td>Allocation</td>
<td>109</td>
<td>5400</td>
<td>600000</td>
<td>72135.41</td>
<td>77363.22</td>
</tr>
<tr>
<td>Resources</td>
<td>109</td>
<td>4</td>
<td>10</td>
<td>6.78</td>
<td>1.49</td>
</tr>
<tr>
<td>Activities</td>
<td>109</td>
<td>2</td>
<td>7</td>
<td>3.58</td>
<td>1.40</td>
</tr>
<tr>
<td>Outputs</td>
<td>109</td>
<td>4</td>
<td>10</td>
<td>6.75</td>
<td>1.38</td>
</tr>
<tr>
<td>Outcomes</td>
<td>109</td>
<td>3</td>
<td>12</td>
<td>6.05</td>
<td>1.90</td>
</tr>
<tr>
<td>Goals</td>
<td>109</td>
<td>4</td>
<td>8</td>
<td>6.06</td>
<td>.91</td>
</tr>
<tr>
<td>Indicators</td>
<td>109</td>
<td>4</td>
<td>16</td>
<td>8.98</td>
<td>3.09</td>
</tr>
<tr>
<td>Evaluation</td>
<td>109</td>
<td>4</td>
<td>13</td>
<td>7.58</td>
<td>2.13</td>
</tr>
<tr>
<td>Logic model</td>
<td>109</td>
<td>20</td>
<td>39</td>
<td>29.22</td>
<td>4.34</td>
</tr>
<tr>
<td>PAQS</td>
<td>109</td>
<td>31</td>
<td>66</td>
<td>46.06</td>
<td>7.70</td>
</tr>
</tbody>
</table>

Research Question One

Does evaluation plan quality significantly relate to organization size for human service nonprofit organizations?

The size of the organization was determined using the Internal Revenue Service Form 990. Tax-exempt, nonprofit organizations with an annual operating budget in excess of $25,000 were required to submit Form 990 each year. A scanned copy of the Form 990 was available on the website: www.guidestar.com. For purposes of this study, the 2005 Form 990s were reviewed. The 2005 forms were selected as this corresponded with the year in which United Way
applications were reviewed to determine funding decisions. Specifically, in order to identify organization size, the total annual operating expenses reported on Form 990 were utilized. After locating the Form 990s that were available on the Internet, there were 17 missing cases. Missing cases included organization such as school districts and other types of entities that received United Way funding yet were not required to submit a Form 990 to the Internal Revenue Service.

The Pearson product-moment coefficient of correlation was calculated to examine the degree of relationship between organization size and evaluation plan quality. Table 2 illustrated the correlation coefficients for each of the PAQS subscales, along with the logic model variable and the PAQS score. An examination of the relationship of organization size with resources, activities, outputs, outcomes, goals, indicators, evaluation, logic model, and PAQS revealed no significant correlations.

Table 2

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>r</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resources</td>
<td>92</td>
<td>-.09</td>
<td>.39</td>
</tr>
<tr>
<td>Activities</td>
<td>92</td>
<td>.07</td>
<td>.53</td>
</tr>
<tr>
<td>Outputs</td>
<td>92</td>
<td>.11</td>
<td>.30</td>
</tr>
<tr>
<td>Outcomes</td>
<td>92</td>
<td>.14</td>
<td>.20</td>
</tr>
<tr>
<td>Goals</td>
<td>92</td>
<td>.12</td>
<td>.26</td>
</tr>
<tr>
<td>Indicators</td>
<td>92</td>
<td>.07</td>
<td>.49</td>
</tr>
<tr>
<td>Evaluation</td>
<td>92</td>
<td>.09</td>
<td>.40</td>
</tr>
<tr>
<td>Logic model</td>
<td>92</td>
<td>.11</td>
<td>.30</td>
</tr>
<tr>
<td>PAQS</td>
<td>92</td>
<td>.12</td>
<td>.27</td>
</tr>
</tbody>
</table>

To further explore the organization size variable, three categories (small, medium and large) for organization size were created. Small (group 1) consisted of an organization size from 0-$499,999, medium (group 2) was organization size between $500,000 and $999,999 and large (group 3) was $1,000,000 to $25,932,720.
Analysis of variance was conducted to determine if there were significant differences between organization size (small, medium or large) and evaluation plan quality (see Table 3). In examining the mean scores, it is important to note that higher scores equates to lower quality plans whereas lower scores equates to higher quality evaluation plans. Findings demonstrated that for each of the variables evaluation plan quality did not significantly differ by organization size.

Table 3

Analysis of Variance Examining Significant Differences Between Organization Size and Evaluation Plan Quality

<table>
<thead>
<tr>
<th>Organization Size</th>
<th>Group 1 (n=32)</th>
<th>Group 2 (n=39)</th>
<th>Group 3 (n=35)</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Resources</td>
<td>6.88</td>
<td>1.26</td>
<td>7.00</td>
<td>1.76</td>
<td>6.42</td>
</tr>
<tr>
<td>Activities</td>
<td>3.40</td>
<td>1.34</td>
<td>3.61</td>
<td>1.50</td>
<td>3.66</td>
</tr>
<tr>
<td>Outputs</td>
<td>6.41</td>
<td>1.58</td>
<td>7.10</td>
<td>1.37</td>
<td>6.69</td>
</tr>
<tr>
<td>Outcomes</td>
<td>6.03</td>
<td>2.06</td>
<td>5.90</td>
<td>1.70</td>
<td>6.31</td>
</tr>
<tr>
<td>Goals</td>
<td>6.03</td>
<td>.90</td>
<td>5.97</td>
<td>.84</td>
<td>6.17</td>
</tr>
<tr>
<td>Indicators</td>
<td>9.06</td>
<td>3.66</td>
<td>8.59</td>
<td>2.84</td>
<td>9.37</td>
</tr>
<tr>
<td>Evaluation</td>
<td>7.63</td>
<td>2.21</td>
<td>7.74</td>
<td>1.87</td>
<td>8.14</td>
</tr>
<tr>
<td>Logic model</td>
<td>28.75</td>
<td>4.48</td>
<td>29.59</td>
<td>4.25</td>
<td>29.26</td>
</tr>
<tr>
<td>PAQS</td>
<td>45.44</td>
<td>8.44</td>
<td>45.92</td>
<td>7.24</td>
<td>46.77</td>
</tr>
</tbody>
</table>

p<.05

Research Question Two

Does evaluation plan quality significantly relate to funding allocation size for human service nonprofit organizations?

The Pearson product-moment coefficient of correlation was calculated to examine the degree of relationship between funding allocation and evaluation plan quality. Table 4 illustrates the correlation coefficients for each of the PAQS subscales along with the PAQS score. There was a significant relationship between funding allocation and the outcomes subscale (p=.02). The
relationship between funding allocation and the outcomes subscale was positive, meaning that as
the funding allocation increased, the score on the outcomes subscale also increased. It is
important to note that although this correlation was positive, it means that as the funding
allocation increased, organizations had lower quality outcomes. There were no significant
relationships between size of funding allocation and each of the remaining subscales (resources,
activities, outputs, goals, indicators and evaluation) and the PAQS score.

Table 4

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>r</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resources</td>
<td>109</td>
<td>-.03</td>
<td>.72</td>
</tr>
<tr>
<td>Activities</td>
<td>109</td>
<td>-.06</td>
<td>.52</td>
</tr>
<tr>
<td>Outputs</td>
<td>109</td>
<td>.01</td>
<td>.88</td>
</tr>
<tr>
<td>Outcomes</td>
<td>109</td>
<td>.23</td>
<td>.02*</td>
</tr>
<tr>
<td>Goals</td>
<td>109</td>
<td>.01</td>
<td>.96</td>
</tr>
<tr>
<td>Indicators</td>
<td>109</td>
<td>.11</td>
<td>.24</td>
</tr>
<tr>
<td>Evaluation</td>
<td>109</td>
<td>.04</td>
<td>.72</td>
</tr>
<tr>
<td>Logic model</td>
<td>109</td>
<td>.08</td>
<td>.44</td>
</tr>
<tr>
<td>PAQS</td>
<td>109</td>
<td>.10</td>
<td>.31</td>
</tr>
</tbody>
</table>

p<.05*

To further explore the funding allocation variable, three categories were created. Group 1
included funding allocations from $5,400 to $49,999, group 2 included funding allocations
between $50,000 and $99,999 and group 3 included funding allocations from $100,000 to
$600,000.

An analysis of variance was conducted to determine if there were significant differences
between the funding allocation groups and the PAQS variables (see Table 5). In examining the
mean scores, it is important to note that higher scores equates to lower quality plans whereas
lower scores equates to higher quality evaluation plans. The activities subscale significantly
differed by funding allocation groups (p=.01). Post hoc analysis using the Bonferroni technique
revealed that group 1 was significantly different from group 3, whereby the group with the smallest funding allocation (group 1) had a lower quality activities subscale than the group with the largest funding allocation (group 3). The outputs subscale significantly differed by funding allocation groups \((p=.04)\). Post hoc analysis using the Bonferroni technique demonstrated that group 1 was significantly different than group 3. Therefore, the group with the smallest funding allocation (group 1) had lower quality outputs subscale than the largest funding allocation group (group 3). There were no significant differences when examining funding allocation groups and the resources, outcomes, goals, indicators, evaluation, logic model or PAQS scores.

Table 5

*Analysis of Variance Examining Significant Differences Between Funding Allocation Groups and Evaluation Plan Quality*

<table>
<thead>
<tr>
<th>Funding Allocation</th>
<th>Group 1 ( (n=52) )</th>
<th>Group 2 ( (n=30) )</th>
<th>Group 3 ( (n=27) )</th>
<th>( F )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resources</td>
<td>6.75 1.31</td>
<td>7.13 1.59</td>
<td>6.44 1.64</td>
<td>1.56</td>
<td>.22</td>
</tr>
<tr>
<td>Activities</td>
<td>4.06 1.35</td>
<td>3.33 1.32</td>
<td>2.93 1.30</td>
<td>7.14</td>
<td>.00*</td>
</tr>
<tr>
<td>Outputs</td>
<td>6.98 1.24</td>
<td>6.87 1.50</td>
<td>6.19 1.39</td>
<td>3.21</td>
<td>.04*</td>
</tr>
<tr>
<td>Outcomes</td>
<td>5.98 1.84</td>
<td>5.90 1.71</td>
<td>6.33 2.22</td>
<td>.43</td>
<td>.66</td>
</tr>
<tr>
<td>Goals</td>
<td>6.17 .81</td>
<td>6.07 .87</td>
<td>5.85 1.10</td>
<td>1.12</td>
<td>.33</td>
</tr>
<tr>
<td>Indicators</td>
<td>8.98 2.94</td>
<td>8.37 2.67</td>
<td>9.67 3.71</td>
<td>1.27</td>
<td>.29</td>
</tr>
<tr>
<td>Evaluation</td>
<td>7.77 2.29</td>
<td>8.10 1.47</td>
<td>7.74 2.46</td>
<td>.28</td>
<td>.76</td>
</tr>
<tr>
<td>Logic model</td>
<td>29.94 3.87</td>
<td>29.30 4.75</td>
<td>27.74 4.52</td>
<td>.23</td>
<td>.10</td>
</tr>
<tr>
<td>PAQS</td>
<td>46.69 7.79</td>
<td>45.77 6.80</td>
<td>45.15 8.60</td>
<td>.38</td>
<td>.68</td>
</tr>
</tbody>
</table>

\(p<.05*\)

*Research Question Three*

Does evaluation plan quality significantly differ by type of human service nonprofit organization?

United Way differentiated human service organization programs into three categories. As represented in Table 6, group 1 \( (n=15) \) were programs that served young children. Group 2
(n=41) were programs that served children and teens. Group 3 (n= 53) were programs that sought to empower vulnerable families. Higher scores on the PAQS equated to lower quality. Analysis of variance revealed significant differences between type of human service organization program in examining the outputs subscales (p=.01). Post hoc analysis using the Bonferroni technique revealed that group 2 was significantly different from group 3. Therefore, group 2, which were organizations that served children and teens, had a lower quality outputs subscale when compared to organizations that empowered vulnerable families (group 3). Additionally, there were significant differences between type of human service organization programs when examining the outcomes subscale (p=.03). Post hoc analysis using the Bonferroni technique revealed that group 1 was significantly different than group 3, whereby organizations that served young children had higher quality outcomes when compared to organizations that empowered vulnerable families. There were no significant differences between type of human service organizations and resources, activities, goals, indicators, evaluation, logic model and PAQS scores.
Table 6

Analysis of Variance Examining Significant Differences Between Type of Human Service Organization and Evaluation Plan Quality

<table>
<thead>
<tr>
<th>Type of Human Service Organization</th>
<th>Group 1 (n=15)</th>
<th>Group 2 (n=41)</th>
<th>Group 3 (n=53)</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Resources</td>
<td>6.60</td>
<td>1.30</td>
<td>6.66</td>
<td>1.65</td>
<td>6.92</td>
</tr>
<tr>
<td>Activities</td>
<td>4.07</td>
<td>1.62</td>
<td>3.71</td>
<td>1.23</td>
<td>3.34</td>
</tr>
<tr>
<td>Outputs</td>
<td>6.93</td>
<td>.88</td>
<td>7.22</td>
<td>1.29</td>
<td>6.34</td>
</tr>
<tr>
<td>Outcomes</td>
<td>5.07</td>
<td>1.71</td>
<td>5.88</td>
<td>1.82</td>
<td>6.45</td>
</tr>
<tr>
<td>Goals</td>
<td>6.20</td>
<td>.77</td>
<td>6.00</td>
<td>.97</td>
<td>6.08</td>
</tr>
<tr>
<td>Indicators</td>
<td>7.87</td>
<td>3.66</td>
<td>8.85</td>
<td>2.38</td>
<td>9.40</td>
</tr>
<tr>
<td>Evaluation</td>
<td>7.47</td>
<td>2.77</td>
<td>7.63</td>
<td>1.81</td>
<td>8.13</td>
</tr>
<tr>
<td>Logic model</td>
<td>28.87</td>
<td>4.69</td>
<td>29.46</td>
<td>3.96</td>
<td>29.13</td>
</tr>
<tr>
<td>PAQS</td>
<td>44.20</td>
<td>10.50</td>
<td>45.95</td>
<td>6.29</td>
<td>46.66</td>
</tr>
</tbody>
</table>

p<.05*

Summary

The purpose of this study was to examine the relationship between evaluation plan quality and organization size, size of funding allocation and type of human service organization. Findings from the study revealed no significant relationships between evaluation plan quality and organization size. There were limited significant findings in the relationship between evaluation plan quality and size of funding allocation although the total PAQS score did not reveal significant findings between evaluation plan quality and size of funding allocation. There were also limited significant finding between evaluation plan quality and the type of human service organizations.
CHAPTER V. DISCUSSION AND RECOMMENDATIONS

Introduction

This study focused on the relationship between evaluation plan quality and factors such as organization size, funding allocation and the type of human service organization. This research utilized the Program Accountability Quality Scale (PAQS) to measure evaluation plan quality. Comprised of seven subscales, the PAQS was created to explore evaluation plan quality within United Way funded organizations. This study utilized the PAQS instrument to assess evaluation plans and the relationship between evaluation plan quality and several organization factors were explored. This chapter contains an overview of the study, discussion, recommendations, future research and conclusions.

Overview of the Study

This study focused on human service nonprofit organizations which were defined as organizations that had obtained 501(c)(3) status through the federal government and were created to serve a charitable purpose that focused on meeting the physical or psychological needs of individuals. Human service nonprofit organizations comprise an important aspect of society as they are often responsible for fulfilling a range of basic needs such as food, shelter, clothing, mental health services, education, and advocacy (Czerwinski, 2007). Within the nonprofit sector, the role of accountability and evaluation has become more important. While a financial reporting of expenditures is expected by the government, other funding sources such as individual donors, foundations, and the public have encouraged the use of programmatic outcomes for nonprofit organizations (Cutt & Murray, 2000; Murray, 2001; Poole et al., 2001; Thayer & Fine, 2001).

Programmatic outcomes include measures that attempt to document the changes produced by participating in or receiving services from a specific program (Connolly & York,
Documenting program outcomes was intended to provide a more substantial and meaningful measure that demonstrated the ability of the organizations to meet their mission (Connolly & York, 2002; Cutt & Murray, 2000).

The establishment of nonprofit evaluation systems has been influenced by evaluation efforts of the for profit, business sector (Bielefeld, 2006; Chetkovich & Frumkin, 2003). However, nonprofit leaders often point out that there are critical differences between the missions, operations, and evaluation goals of nonprofit versus for profit entities. Nonprofit evaluation systems are often more complex as there are varying stakeholders that may have differing perspectives on the role and purposes of the services provided by the organization (Krug & Weinberg, 2004; McHargue, 2003; Wholey et al., 1986).

The increased focus on accountability within the nonprofit sector has resulted in numerous evaluation approaches (Bozzo, 2000; Julian & Kombarakaran, 2006; Murray, 2001). Research focused on evaluation efforts, however little attention has been given to evaluation plan quality. Evaluation plan quality is of critical importance as it provides a foundation for the evaluation process. Poole, et al. (2000) developed the Program Accountability Quality Scale (PAQS) as a method to assess the quality of evaluation plans. The PAQS contains seven subscales that relate to program service areas specifically identified for human service nonprofit organizations (resources, activities, outputs, outcomes, goals, indicators, and evaluation).

This study intended to further the limited research that exists in the area of evaluation plan quality and explored factors related to evaluation plan quality in human service nonprofit organizations. A correlational research design was used to explore the relationship between evaluation plan quality and three organization factors (organization size, funding allocation size and type of organization). The PAQS instrument was utilized to assess the level of evaluation
plan quality. The sample consisted of 125 human service nonprofit organizations that received funding from a United Way organization in Ohio in 2006.

Discussion

Discussion of Descriptive Results.

Descriptive results showed that the range of organization size was from $112,999 to $25,932,720. These results demonstrated a wide range of operating budgets among the organizations funded by United Way. Likewise, there were a wide range of funding allocations from $5,400 to $600,000, which revealed that the level of funding varied, based on the nature and intensity of the program and the amount requested by the organization.

Discussion of Organization Size and Evaluation Plan Quality.

The first research question examined if evaluation plan quality significantly differed by organization size for human service nonprofit organizations. These factors were examined by utilizing the Pearson product-moment coefficient of correlation to determine if there was a significant relationship between organization size and the PAQS. There were no significant relationships between organization size and the seven PAQS subscales. Additionally, there were no significant relationships between organization size and the logic model score or the total PAQS score. This analysis revealed that the level of evaluation plan quality was not related to organization size.

An analysis of variance was also completed that explored the relationship between organization size and evaluation plan quality. For purposes of this analysis, organization size was divided into three categories (small, medium, large). The ANOVA results demonstrated that there were no significant differences between organization group size and evaluation plan quality.
Findings from the correlation and ANOVA statistics demonstrated that evaluation plan quality did not significantly relate to organization size or differ by organization size groups. This research question was of interest, due to previous research that examined the relationship between organization size and evaluation and program outcome quality. Based on the results of this data analysis, organization size did not significantly relate to evaluation plan quality. This finding relates to the research presented by Robertson (2007) that found that large organizations (school districts) did not have a significant impact on program outcomes. Although large organizations may have access to additional and more advanced resources, the availability of financial resources within an organization did not influence evaluation plan quality. Similar to the rationale for the findings presented by Robertson (2007), the larger organizations in the current study might not have demonstrated higher quality measures because, due to their size and bureaucratic processes, they are unable to be as flexible and adaptive.

While the finding that organization size did not significantly relate to evaluation plan quality is related to the findings by Robertson (2007), it does not support the research findings reported by B. R. Lee et al. (2007). B. R. Lee et al. found that particular aspects of quality initiatives significantly related to organization size. Generally, larger organizations engaged in more quality related initiatives when compared to the smallest organizations in the sample. However, it is important to note that organization size was measured in different ways between the two studies. The current study utilized annual operating revenues as the measure of organization size and the study completed by Lee et al. utilized number of employees. Both of these measures are acknowledged as appropriate aspects of organization size (Ebaugh et al., 2005). Even with the difference in organization size variables between the two studies, the results highlight an area of future exploration.
In this study, organization size was defined by the organization’s annual operating expenses. This measure of organization size was identified based on the consistency of information submitted on the Internal Revenue Service Form 990. Annual operating expenses have been acknowledged as an appropriate measure of organization size by other researchers (Ebaugh et al., 2005; B. R. Lee et al., 2007). Future research could explore the relationship between alternate organization size variables such as number of employees or number of programs/activities and evaluation plan quality.

Research findings presented by Julian and Kombarakaran (2006) demonstrated that when United Way implemented an outcome evaluation measurement system that the evaluation plan quality of all United Way funded organizations increased. Therefore, this finding seems to be in unison with the findings presented by Julian and Kombarakaran (2006) as the implementation of the outcome evaluation system by this United Way may have had a positive impact on the evaluation plan quality of all funded organizations.

Discussion of Funding Allocation and Evaluation Plan Quality.

The second research question examined if evaluation plan quality significantly related to funding allocation size for human service nonprofit organizations. These factors were examined by using the Pearson product-moment coefficient of correlation to determine if there was a significant relationship between funding allocation and evaluation plan quality. Results from the analysis revealed that there was a significant relationship between funding allocation and the outcomes subscale. The outcomes subscale relates to the measurable benefits for program participants and the measurable accomplishments of the program. Organizations that received a larger funding allocation scored higher on the outcomes subscale, which equated to lower quality outcomes. This finding revealed that as funding allocation size increased, the organizations
demonstrated lower quality outcomes. Generally, it would be expected that outcome quality increase (or at a minimum) remain the same as funding allocation size increased. While the finding from this study seems contradictory to conventional thought, it is important to emphasize that the outcomes subscale is a complex subscale. Outcome measurement is considered one of the more challenging aspects of program evaluation as it involves measuring the impact of the program on the person’s well being or health (Bielefeld, 2006; Plantz et al., 1997). Different than other subscales that require the collection and analysis of more basic information, the outcomes subscale measures program impact, an aspect of evaluation that can be difficult to quantify (Plantz et al., 1997). This finding relates to the study completed by Poole et al. (2001) that explored the relationship between organization characteristics and evaluation plan quality. Poole et al. found that lower quality evaluation was related to program outcomes when compared to other PAQS subscales. It was projected that this result was due to the complexity of outcome measurement. While the finding from Poole et al. differs from the current finding, it does bring into question the impact of outcome measurement complexity when assessing evaluation plan quality.

The statistical analysis did not reveal any other significant relationships. Therefore, the size of funding allocation was not related to evaluation plan quality on the resources, activities, outcomes, goals, indicators, evaluation subscale, the logic model or the combined PAQS score.

An analysis of variance was also completed to explore group differences of funding allocation in evaluation plan quality. Funding allocation was divided into three categories to complete the analysis of variance. The ANOVA results revealed two significant relationships. There was a significant relationship between funding allocation groups and the activities subscale. The activities subscale assessed the number and type of activities undertaken by the
program to fulfill its mission. The group of organizations that had smaller funding allocations had lower quality activities when compared to the group of organizations that had the largest funding allocations. There was also a significant relationship between funding allocation groups and the outputs subscale. The outputs subscale focused on the direct products of the program activities. The group of organizations that had smaller funding allocations had lower quality outputs when compared to the group of organizations that had the largest funding allocations. Based on these findings, it appears that there were two subscales in which group funding allocation size had a significant relationship with evaluation plan quality and the group of organizations that had received larger funding allocations achieved a higher quality rating.

Smaller funding allocations are typically distributed to organizations that operate programs that are more limited in scope and intensity than programs that receive medium or large funding allocations. Therefore, the lower quality evaluation plans submitted by the small funding allocation group might be directly related to offering a program that has fewer activities than the large funding allocation group and also might have limited outputs based on the scope of the program due to limited funding. From another perspective, these findings might be related to the situation that organizations that received a large funding allocation might have placed a higher priority on completion of the evaluation plan and this could have resulted in higher quality evaluation plans on the two subscales. Regardless of the size of the funding allocation, to ensure evaluation plan quality it seems as though organizations will have to dedicate resources to the evaluation process. Although resources are scarce and limited for nonprofits, evaluation is increasingly becoming a requirement of grant funding. Therefore, organizations should anticipate allocating a sufficient portion of their resources to the evaluation process and in particular to establishing a quality evaluation plan.
It is interesting to note that statistically significant findings differed between the correlational analysis and the ANOVA. While funding allocation size and the outcomes subscale were significant in the correlational analysis, the relationship was not significant when funding allocation size was distributed into three groups (small, medium and large). Several factors might have contributed to these findings. Considering the overall size of the sample and of the funding allocation groups, it could be that the groups were not large enough to produce consistent results.

Findings from these statistical analyzes revealed that for the majority of the variables, funding allocation did not significantly relate to evaluation plan quality. Therefore, whether an organization received a small, medium, or large funding allocation, there was not a significant relationship with evaluation plan quality. This finding might be due to the situation that organizations have achieved a certain level of evaluation plan quality, regardless of the size of funding allocation.

However, these findings differ from those presented by Brown and Troutt (2002), Carrilio et al. (2003) and B. R. Lee et al. (2007). Brown and Troutt stated that funding restrictions based on size or type of funding negatively impact evaluation, program outcomes and stability of the organization. The negative impact on evaluation was found to be related to restricted funding. The current findings do not correspond with Brown and Troutt’s findings as the size of funding allocation did not significantly relate to evaluation plan quality for each of the variables (other than the outcomes subscale). Findings from Carrilio et al. stated that the utilization of a performance management evaluation system was related to access to sufficient financial resources and stakeholder involvement. However, considering the findings presented by Carrilio, it appears that in the current study financial resources (measured by the size of funding allocation) did not have an impact on evaluation plan quality. Additionally, research findings
presented by B. R. Lee et al. (2007) explained that limited resources were related to quality in terms of program and service offering and evaluation. Although the current study did not explore the quality of program and service offerings, it did examine one particular aspect of evaluation, the evaluation plan. Findings from the current study did not find that limited funding resources impacted evaluation plan quality.

Discrepancies in these findings might be due to the highly structured method by which the United Way allocates funding and assesses outcome measurement. Different from other funders that may accept a range of proposal formats and outcome measurement models, United Way has created a structured system in which organizations complete similar forms and are exposed to similar training sessions regarding the outcome measurement system. It seems as though the lack of significant findings might be due to participation in a structured system and it identifies an area of further exploration to determine if the level of evaluation plan quality is related to the structure of the evaluation process.

Discussion of Type of Organization and Evaluation Plan Quality.

The third research question examined if evaluation plan quality significantly differed by type of human service organization. There were three types of human service organizations that comprised this variable: 1) programs that served young children, 2) programs that served children and teens and 3) programs that intended to empower vulnerable families. Recognizing that organizations with distinctly different mission often have different funding sources, national affiliations, professional associations and operating standards, it was important to examine if there was a relationship between evaluation plan quality and the type of human service organization (Poole et al., 2001).
These factors were examined by using an analysis of variance. The ANOVA results revealed that there were significant differences on the outputs subscale among the three types of human service organizations. The group of organizations that served children and teens scored lower on the outputs subscale when compared to the group of organizations that sought to empower vulnerable families. Additionally, the ANOVA results revealed that there were significant differences on the outcomes subscale among the three types of human service organizations. The group of organizations that served young children scored higher on the outcomes subscale when compared to the group of organizations that sought to empower vulnerable families. Based on these findings, it seems feasible that organizations that have similar mission and purpose might have access to or have participated in training that differs from other types of organizations. For example, the finding that organizations that served young children had higher outcomes subscale quality than organizations that empowered vulnerable families is supported by the movement within organizations that serve young children to document program outcomes. Therefore, lower or higher quality results on the subscales may be due to other resources and opportunities available to nonprofit leaders of these organizations. These findings are similar to the study completed by B. R. Lee et al. that found significant differences between the types of human service organizations on two quality-related measures (quality personnel positions and the establishment of quality processes). In the study completed by B. R. Lee et al., health organizations had higher quality personnel positions and quality processes when compared with organizations that were identified in the other category. The remaining two categories were mental health organizations and health organizations. Although the types of human service organizations identified by B. R. Lee et al. differed from the categories of human service organizations that were utilized in the current study, the findings of
significant differences in both of the studies highlights the need for further research to
understand the possible impact and interaction of type of human service organizations with
efforts related to evaluation quality.

For the remaining subscales (resources, activities, goals, indicators, evaluation) and the
logic model and total PAQS score, there were no significant differences among the types of
human service organizations. Therefore, evaluation plan quality did not relate to type of human
service organization for these variables.

Recommendations

This section details recommendations for funders and nonprofit leaders, based on the
findings from the research study. These recommendations have been designed to inform future
policy and practice among the funding community that includes United Way organizations,
foundations, and nonprofit leaders that are engaged in the evaluation plan process. Nonprofit
leaders can include positions such as executive directors, board members, development directors,
direct service providers and administrative staff, depending on how evaluation responsibilities
are allocated within an organization.

Recommendations Based on Organization Size Findings.

The primary recommendation for funders is that training and individualized technical
assistance is offered to all organizations that receive funding, regardless of the organization size
or the level of funding allocation. Findings from this study demonstrated that evaluation plan
quality, as determined by utilizing the PAQS instrument, was not significantly related to
organization size or funding allocation. Therefore, assistance based on factors such as
organization size, funding allocation size or type of organization may not prove beneficial to
increasing evaluation plan quality. Funders might choose to utilize the PAQS instrument to
determine which organizations have lower quality evaluation plans and then training and individualized technical assistance could be offered to those organizations. This recommendation based on the research findings from this study corresponds with research findings from Broom and Jackson (1997) and Poole et. al. (2001) that recommend evaluation training be tailored to the skills and needs of the organization’s staff members. Acknowledging the limited resources in the nonprofit sector (both for nonprofit organizations and funders) and the amount of time currently dedicated to evaluation efforts (Stoecker, 2007), it seems important that organizations should be specifically targeted for assistance to increase their evaluation plan quality, rather than requiring that all organizations participate in training and technical assistance.

Another recommendation based on the findings from this study and supported by the existing literature is that the Program Accountability Quality Scale or other such instrument could be utilized as an evaluability assessment that would be completed prior to finalizing the evaluation plan (Poole et al., 2001; Weiss, 1998). The PAQS instrument has been established as a method to assess evaluation plan quality. It is recommended that evaluation plans be submitted to the funding entity prior to the organization engaging in the evaluation process. This would allow an opportunity for an assessment of the evaluation plan. Considering that the PAQS instrument contains seven subscales, areas of strengths and deficiencies within the evaluation plan could be efficiently identified and communicated to the organization. Therefore, modifications to the evaluation plan could take place prior to the organization receiving funding for the program and prior to the implementation of the evaluation process. Establishing an evaluation process that allows the PAQS instrument to be used as an evaluability assessment is intended to have several benefits. Primarily, it is intended to increase evaluation plan quality. Beyond increasing evaluation plan quality, it is also intended to create opportunities for
increased communication regarding evaluation expectations between funders such as a United Way organization and nonprofit organizations that are receiving funding. Utilization of the PAQS also allows for a specific feedback from funders regarding the assessment of evaluation plan quality. Additionally, assessing evaluation plan quality may provide nonprofit organizations with an opportunity to analyze and reflect on the evaluation plan process. Engaging in this type of process, prior to program implementation, may create opportunities for multiple staff members at an organization to be involved in the evaluation process.

Another aspect of organization factors is the role of stakeholder involvement. Previous research demonstrated the critical role of stakeholders in the perceptions of the use and value of evaluation within the organization (Cutt & Murray, 2000; Fine et al., 2000; Newcomer, 1997; Wholey & Newcomer, 1997). Based on the research findings and the findings from the current study, further examination of the role of stakeholders in the evaluation process might provide additional insight into evaluation plan quality. The role of stakeholders in evaluation plan quality and the overall evaluation process might be ascertained through a qualitative study of the value and role of the evaluation process within the organization.

**Recommendations Based on Funding Allocation Findings.**

Results from the analysis of the relationship between evaluation plan quality and funding allocation revealed a significant relationship between the outcomes subscale and funding allocation. Although the finding is contradictory to conventional thought, it possibly highlights an area of deficiency within the nonprofit organizations that were included in this study. It is recommended that training and technical assistance specifically on outcome measurement be offered on an ongoing basis to nonprofit organizations. It may also be reasonable that funders,
such as the United Way in this study, establish particular evaluation plan quality benchmarks that organizations need to achieve before initiating the program.

Based on the significant findings that the activities subscale and the outputs subscale significantly related to evaluation plan quality, it is recommended that the outcome measurement model be examined. Although organizations are required to complete a similar outcome measurement plan, regardless of the size of funding allocation, the findings from this study highlight questions regarding the appropriateness of requiring a similar system for all organizations. Recognizing the impact of the funding allocation size on the intensity and scope of a particular program, it may be unrealistic to expect that organizations that receive a small funding allocation will obtain a similar level of evaluation plan quality when compared to organizations that receive a medium or large funding allocation. Considering that one of the desired uses of evaluation that has not yet been achieved is to inform funding decisions (Campbell, 2002; Cutt & Murray, 2000; Wholey & Newcomer, 1997), the uniform expectations in regard to outcome measurement, regardless of funding allocation should be thoroughly explored.

Recommendations Based on Type of Human Service Organization Findings.

Findings from the third research question that explored the relationship between evaluation plan quality and type of human service organization revealed some interesting results. Although the results were not consistent, there was variation among the quality of the outputs and outcomes subscales among the types of human service organizations. These findings might be due to the impact that other resources, beyond the United Way, had on evaluation plan quality. These resources might include opportunities such as professional networking, national publications, discipline specific training and education. Organizations should be encouraged to
utilize resources such as this to increase their organizational capacity in regard to evaluation. Additionally, United Way and other funders should consider offering training sessions specifically designed for the various types of human service organizations that they fund. Trainings that are offered to similar organizations would provide an opportunity to discuss areas of concerns specific to those organizations. Additionally, the training could be tailored to provide specific examples and information based on the type of human service organization.

*Future Research*

While this study contributed to the literature base regarding the role and use of evaluation efforts in the nonprofit sector, other research studies could be completed to further knowledge in regard to this topic. As this study focused on assessing evaluation plan quality, a subsequent study could explore the relationship between evaluation plan quality and evaluation outcomes. For example, it would be insightful to explore if organizations that had higher quality evaluation plans had more successful program outcomes and likewise, if organizations that had lower quality evaluation plans had less successful program outcomes. A study of this nature would provide insight into the role of evaluation preparation and planning prior to the establishment of human service programs. Additionally, a study of this nature could explore the implementation of the evaluation plan and the relationship between evaluation plan implementation and program evaluation outcomes.

Another research study that could be undertaken to further the literature could focus on assessing evaluation plan quality among programs funded by several United Way organizations. One of the limitations of the current study was that the sample consisted of organizations that were funded by the same United Way organization. Due to the situation that the current study focused on one United Way organization, it was not feasible to determine the impact that the
United Way had on evaluation plan quality. United Ways typically offer training sessions and at times, individualized technical assistance to assist organizations as they complete their evaluation plans. Training sessions are typically offered prior to evaluation plan submission while individualized technical assistance can occur throughout the allocation process. The individualized technical assistance can also take place after the submission of the evaluation outcomes, which typically occur at six month, one year, two year and three-year periods. A future research study could explore differences in regard to evaluation between the various United Way organizations with a specific emphasis on exploring the impact of the intensity and type of United Way relationship and evaluation plan quality. A study such as this could explore if there is a relationship between the level of training that organizations participate in and evaluation plan quality. Additionally, the specific type of training that organizations participate in and the relationship with evaluation plan quality could be examined.

A qualitative case study exploration of the relationship between a human service nonprofit organization, the local United Way and other funding entities could also further knowledge on the relationship between funding entities and evaluation plan quality. A study of this nature could focus on understanding the evaluation plan and outcome process, which may provide further insight into the dynamics of creating an evaluation plan. A case study such as this could provide insight into the factors such as organizational culture, management support, technical capabilities, stakeholder involvement (Poole et al., 2001) along with the relationship with funders and the role of evaluation within the organization. The case study methodology proposed for future research could provide much needed insight into the evaluation plan process.

Furthermore, the current study was exploratory in nature and therefore, utilization of the PAQS instrument to assess evaluation plan quality in relation to other organization factors
should be examined. The current study revealed some significant findings; however, these should be interpreted with caution. In regard to future research specifically with the PAQS instrument, the validity and strength of the PAQS instrument could be explored. Assessing the strength of the instrument could provide further insight into ways in which the PAQS might be utilized either by individual organizations or by funding entities in the future. Although the PAQS instrument was specifically created to assess evaluation plan quality of United Way funded organizations, future research could explore the applicability of the PAQS instrument to other United Way chapters and other funders beyond United Way.

Conclusions

Four conclusions can be drawn from this study that examined the relationship between evaluation plan quality and organization factors: 1) the size of the human service organization was not related to evaluation plan quality, 2) while the results of the relationship between funding allocation and evaluation plan quality varied, it did appear that on a few of the measures organizations that received less funding exhibited lower quality evaluation plans, 3) there were some significant differences between some types of human service organizations and evaluation plan quality and 4) there is a critical need for further research to understand evaluation systems utilized by nonprofit organizations.

The research findings demonstrated that evaluation plan quality did not relate to organization size. Therefore, although there were differences in relation to financial resources between the organizations it did not impact evaluation plan quality. This finding is supported by research conducted by Julian and Kombarakaran (2006) that found that implementation of an outcome evaluation system initiated by a funder had a positive impact on the evaluation plan quality for all of the organizations that received funding. It appears that the establishment of a
formal system by United Way along with the training and technical assistance that is provided may have contributed to the lack of significance regarding the organization size variable.

In regard to the relationship between size of funding allocation and evaluation plan quality, the activities and outputs subscales resulted in lower quality scores for organizations that received smaller funding allocations when compared to those organizations that received the largest funding allocations. These findings bring into question the applicability of requiring the same outcome measurement system regardless of funding allocation size.

This study demonstrated some significant differences between type of human service organization and evaluation plan quality. These findings highlight the resources and training that may be provided to certain segments of the human service sector that impacts the quality of the evaluation process. As this is an area of limited research, further study would need to be undertaken to understand the impact of type of human service organization and the relationship with evaluation plan quality.

Although evaluation in the nonprofit sector is considered to be in its infancy and has not always been well received by nonprofit leaders (Bielefeld, 2006), the existing literature base documents the growing importance and usage of evaluation within the sector. While one of the primary driving forces behind the evaluation and accountability movement in the nonprofit sector has been for funders to utilize the evaluation results to determine future funding for programs or projects (Wholey & Newcomer, 1997), this goal cannot be achieved until quality evaluation plans have been established. Therefore, future research efforts to deepen the understanding of evaluation will be critical in funding decisions.

Furthermore, although human service organizations in the nonprofit sector have at times considered their uniqueness to be a barrier to implementing effective evaluation systems, this
sentiment was shared by the business sector decades ago (Beck et al., 2008). When total quality management and other quality control techniques were first introduced, businesses resisted implementation of the practices stating that the business sector was too diverse to implement such practices. Just as the business sector successfully adopted new management approaches, the growing base of research literature and the commitment among nonprofit leaders will propel the nonprofit sector to advance the quality and utilization of evaluation.
REFERENCES


PROGRAM ACCOUNTABILITY QUALITY SCALE (PAQS)
© Heart of Florida United Way

Agency: Dr. Nelson Ying Center
1940 Traylor Blvd.
Orlando FL 32804

Program:

Rating Scale
1 = Strongly Agree/Very Good Performance–only minor revisions, if any.
2 = Agree/Good Performance–some strengths, some areas need revision.
3 = Disagree/Fair Performance–few strengths, major revisions required.
4 = Strongly Disagree/Poor Performance–lacking strengths, insufficient information provided.

LOGIC MODEL

Definitions
Resources: The program’s essential ingredients. Resources dedicated to or consumed by the program.
Activities: What the program does with the resources to fulfill its mission. (How does staff spend their time?)
Outputs: The direct products of the program activities. (How many people does the program serve? For what period of time?)
Outcomes: Benefits for program participants. What is the program trying to accomplish?
Goals: The MISSION. What is the larger impact?

Resources:
Are the following resources listed (not scored, use to answer questions 1 – 3)?:

<table>
<thead>
<tr>
<th>Resources</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service Providers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Program Setting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community Factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collaborations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Service Technologies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Funding Sources</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Most areas of resources are addressed. 1 2 3 4
2. The resources seem comprehensive. 1 2 3 4
3. The resources seem to match this type of program. 1 2 3 4

Comments:

Total Score:
### Activities:

4. The activities logically link to the outputs listed. 1 2 3 4
5. There are sufficient activities to achieve the outcomes. 1 2 3 4

Comments:

<table>
<thead>
<tr>
<th>Total Score:</th>
</tr>
</thead>
</table>

### Outputs:

6. The numbers of participants are identified for each activity. 1 2 3 4
7. The numbers of events/process are listed (*dose*). 1 2 3 4
8. Time frames are given for outputs (*duration*). 1 2 3 4

Comments:

<table>
<thead>
<tr>
<th>Total Score:</th>
</tr>
</thead>
</table>

### Outcomes:

9. The outcomes logically link to the goal(s). 1 2 3 4
10. The outcomes are written as change statements. 1 2 3 4
11. The outcomes are truly outcomes rather than activities or outputs. 1 2 3 4

Comments:

<table>
<thead>
<tr>
<th>Total Score:</th>
</tr>
</thead>
</table>
Goal(s):
12. The program goals indicate the intended effect of the program on the need of the population. 1 2 3 4
13. The program goals describe the broad community impact. 1 2 3 4

Comments:

Total Score:

INDICATORS

14. The indicators are stated in specific and measurable terms. 1 2 3 4
15. The indicators are valid measures of outcomes. 1 2 3 4
16. The indicators will efficiently measure progress toward achievement of the outcomes. 1 2 3 4
17. The indicators are important to the changes program planners want to measure. 1 2 3 4

Comments:

Total Score:

EVALUATION PLANS

18. The data collection method will generate reliable information. 1 2 3 4
19. The evaluation plan can be implemented with available resources. 1 2 3 4
20. The evaluation plan is designed to measure progress toward outcomes in an efficient manner. 1 2 3 4
21. The evaluation plan is realistic. 1 2 3 4

Comments:

Total Score:
# Outcome Measurement System Summary

<table>
<thead>
<tr>
<th>Section</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resources</td>
<td></td>
</tr>
<tr>
<td>Activities</td>
<td></td>
</tr>
<tr>
<td>Outputs</td>
<td></td>
</tr>
<tr>
<td>Outcomes</td>
<td></td>
</tr>
<tr>
<td>Goals</td>
<td></td>
</tr>
<tr>
<td>Logic Model Score (a)</td>
<td></td>
</tr>
<tr>
<td>Indicators Score (b)</td>
<td></td>
</tr>
<tr>
<td>Evaluation Score (c)</td>
<td></td>
</tr>
<tr>
<td><strong>Total Score (a+b+c)</strong></td>
<td></td>
</tr>
</tbody>
</table>

**Summary Comments:**