2013

Program Evaluation in the Nonprofit Sector : An Exploratory Study of Leaders' Perceptions

Gail A. Patin
University of North Florida, gap350@hotmail.com

PROGRAM EVALUATION IN THE NONPROFIT SECTOR:

AN EXPLORATORY STUDY OF LEADERS' PERCEPTIONS

By

Gail A. Patin

A dissertation submitted to the Doctoral Faculty of the College of Education

and Human Services in Educational Leadership in partial fulfillment for the degree of

Doctor of Education

UNIVERSITY OF NORTH FLORIDA

COLLEGE OF EDUCATION AND HUMAN SERVICES

August, 2013

The dissertation of Gail Patin is approved:

Katherine L. Kasten, Chair Date

Cheryl Fountain, Ed.D. Date

Warren Hodge, Ph.D. Date

Georgette Dumont, Ph.D. Date

Accepting for the Department:

Jennifer J. Kane, Ph.D., Chair Date

Department of Leadership, School Counseling & Sport Management

Accepting for the College:

Larry G. Daniel, Ph.D., Dean Date

College of Education & Human Services

Accepting for the University:

Len Roberson, Ph.D., Dean Date

The Graduate School

Acknowledgements

There are not enough words or ways to thank the many people who have traveled with me on this journey. First, I truly thank Dr. Katherine Kasten for serving as my dissertation chair. Her wisdom, guidance, and patience helped me stay the course. I am a better person for knowing her.

I must also thank the rest of my dissertation committee: To Dr. Warren Hodge for his recommendations that helped bring out the best in me; to Dr. Cheryl Fountain for her astuteness and keen attention to detail; to Dr. Georgette Dumont for her knowledge and zeal regarding the nonprofit sector. I was truly blessed to be in the company of such remarkable people.

I must also thank the leaders in the Northeast Florida nonprofit sector that participated in the study. Their interest in the subject, gift of time, and wisdom are deeply appreciated.

I must thank my supervisor and friend, Ellen Siler, for all of the support and flexibility she afforded me during the course of this venture. She made a dream possible.

I also must thank Dr. Walter Shepherd for helping me know myself better and Cohort 18 for sharing this adventure with me.

I owe a huge debt of gratitude to my family and friends. Seeking a doctorate degree is ultimately a very selfish endeavor. I cannot thank my family and friends enough for giving me the space to achieve this goal. Their understanding of my many absences from their lives has not gone unnoticed. Thank you to my Mom and Marina for having confidence in me. Thank you to my Father for his example of the pursuit of higher education.

Finally, I would like to thank Kim Marsh for the many Saturdays and Sundays she spent alone while I studied. Most importantly, her unfailing love and humor have sustained me.

Table of Contents

List of Tables

List of Figures

Abstract

The purpose of the study was to understand the role of and capacity for program evaluation from the perspectives of leaders in the Northeast Florida nonprofit sector. The study was a two-phased qualitative design and had three research questions related to the role of program evaluation, the capacity for program evaluation, and the influence of the relationship between providers and funders on these efforts. The first phase was a nominal group process using the Delphi survey method. The second phase involved in-depth interviews. Study participants were chosen based on specific criteria.

The Delphi phase had four rounds of surveys. Findings from the Delphi phase of the study indicated that leaders in Northeast Florida perceived there were eight distinct roles of program evaluation. The roles included assessing impact of programs, advancing organizational learning, cultivating funding collaborations, informing program management decisions, enhancing communication with multiple stakeholders, facilitating quality assurance, determining resource allocation, and validating organizational credibility. Findings regarding the essential capacity elements needed for program evaluation efforts included having sufficient time, sufficient financial resources, a positive culture, functional program evaluation designs/methods, sufficient human resources, realistic expectations from the philanthropic community, ongoing collaboration, and ongoing training. Study participants reported that these capacity elements were lacking in the Northeast Florida nonprofit sector.

Interviews produced similar findings as the Delphi survey. In particular, the use of reflective practice as a role within program evaluation efforts and as a program evaluation approach were core topics of interviews. Additional findings from the study were related to program evaluation capacity development strategies, the dynamics of the relationships between

providers and funders, and the influence of these variables on the culture of the Northeast Florida nonprofit sector.

Chapter 1: Introduction

The nonprofit sector is an integral part of American society. Nearly 1.5 million nonprofit organizations are registered with the IRS, representing almost $2 trillion of revenue and over $4 trillion in assets (Roeger, Blackwood, & Pettijohn, 2011). Increasingly, government is divesting itself from providing essential services to United States citizens and contracting with nonprofits to fill the void (Candler & Dumont, 2010; Hall, 2010). The evolution of the nonprofit sector from a grassroots movement to a multi-million dollar industry has prompted the public, funders, and regulators to require accountability from nonprofits through assessment of organizational and program effectiveness. In short, communities want to know that their dollars invested produce results and make an impact in society as commissioned.

Although the genesis of evaluating program effectiveness harkens back to the 1960s, the push for organizational accountability through determination of effectiveness in the nonprofit sector took root in the 1990s. Furthermore, funding sources increasingly are linking documenting outcomes and impact to continuation of funds (Hendricks, Plantz, & Pritchard, 2008). Currently, the intense focus on validating the efficacy of nonprofit programs does not appear to be decreasing.

Unfortunately, nonprofits are playing catch-up to meet the growing accountability demands for information regarding effectiveness that results from quality program evaluation efforts. Funders and nonprofits allocate minimal resources for program evaluation. Nonprofit staff rarely have the time, resources, education, skills, or knowledge to conduct quality program evaluations (Carman, 2007; Carman & Fredericks, 2009). Furthermore, staffing issues of nonprofits also inhibit institution of a quality program evaluation agenda (Kegeles, Rebchook, & Tebbetts, 2005). This lack of program evaluation capacity has direct bearing on the quality of program evaluation efforts in the nonprofit sector.

Even so, information regarding the use and sustainability of quality program evaluation agendas in the nonprofit sector is unclear. In particular, the level of collaboration between leaders of nonprofit organizations and funding organizations influence the uses of and capacity for program evaluation in the nonprofit sector (Carman & Fredericks, 2009). The lack of candidness between these two groups of leaders seems embedded in the type of relationships forged. In a recent study, Carman (2011) found most of the nonprofits in the study participated in program evaluation efforts only at the bequest of funders or because of other external pressures. Participation in program evaluation was only for compliance purposes and not to increase the effectiveness of the program. Sixteen percent used program evaluation as a means to an end such as to secure more resources. Only 23% of the agencies reported using program evaluation to inform decision-making and improve performance. Clearly, the reasons for implementing program evaluation are diverse, and the majority of the nonprofits examined did not initiate program evaluation efforts for their constituents' benefit or to promote organizational learning (Carman, 2011).

Statement of Purpose

As a leader in the nonprofit sector for over 25 years, I have experienced the growing mandates from funders and other stakeholders for documenting successful outcomes from the programs I have supervised. Juxtaposed with these external pressures for accountability was my own internal need to know if my agency was making a positive difference in the lives of our constituents. However, I experienced much frustration with these demands due to limited program evaluation capacity issues such as lack of knowledge and resources. Yet I sensed that the current zeitgeist regarding program evaluation in the nonprofit sector was ripe for cooperative program evaluation strategies. More pointedly, I have operated from the assumption that the manner of collaboration between funders and management of nonprofit organizations directly influences program evaluation efforts (Carman, 2011). Consequently, I surmised that learning the perceptions regarding this topic from both human services providers' (providers) and funding organizations' (funders) leaders in the nonprofit sector would be essential.

For the present study, it was important to gain knowledge regarding the role of and capacity for program evaluation from the leaders of nonprofit human services providers and funding organizations, as both groups have influence to shape future program evaluation agendas and policies for the nonprofit sector. Therefore, the purpose of the present study was to understand the role of and capacity for program evaluation in the nonprofit sector from the perspectives of nonprofit leaders.

Significance of the Study

The nonprofit sector is at a critical development point. Gone are the days of volunteers haphazardly putting together programs to help people. As external forces place more services on nonprofits or during a time of economic downsizing, the demand for demonstrating accountability or a return on investments is greater (Behrens & Kelly, 2008; Candler & Dumont, 2010; Murray, 2010; Woodwell, 2005). The present study provided information about the perceptions among the leaders of nonprofit human services providers and funding organizations regarding the role of and capacity for program evaluation in the Northeast Florida nonprofit sector. Nonprofit organizations and funders may use the information to create common understanding of the multiple factors related to the role of program evaluation in the nonprofit sector as well as the elements of program evaluation capacity. Additionally, the present study advanced knowledge in the field by providing information that may help (a) develop stronger collaborations between leaders of funding and nonprofit organizations, (b) advance strategic planning around program evaluation agendas, and (c) increase understanding regarding the essential elements needed for program evaluation capacity in nonprofit organizations to meet program evaluation demands.

Research Questions

The purpose of this present study was to understand the role of and capacity for program evaluation in the Northeast Florida nonprofit sector from the perspectives of leaders from nonprofit human services providers and funding/philanthropic organizations. The three research questions posited for the present study were the following:

- What are the perceptions of nonprofit and funding organizations' leaders regarding the role (e.g., meaning, purposes, benefits, approaches, and utilization) of program evaluation in the Northeast Florida nonprofit sector?
- What are the perceptions of nonprofit and funding organizations' leaders regarding how the relationships between service providers and funders influence program evaluation efforts in the Northeast Florida nonprofit sector?
- What are the perceptions of nonprofit and funding organizations' leaders regarding the development of program evaluation capacity in the Northeast Florida nonprofit sector?

Summary of Methods and Procedures

I used the Delphi method and in-depth interviews to ascertain the perceptions of the leaders from both nonprofit human services providers and funding organizations. The Delphi method uses experts in the field to obtain subjective knowledge and to quantify the collective consciousness of the selected group regarding a specific topic (Geist, 2010; Hung, Altschuld, &

Lee, 2008; Landeta, 2006; Skulmoski, Hartman, & Krahn, 2007). In the present study, the leaders of nonprofit organizations and funding organizations were the experts providing information on the role of and capacity for program evaluation in the Northeast Florida nonprofit sector. I used purposeful sampling to select panelists for the present study. I used multi-round Internet-based surveys to gather the data and provide feedback to participants. As the researcher, I was the only one who knew the identity of participants and individual responses. However, participation among panelists was anonymous, which is consistent with recommendations from the literature on the implementation of the Delphi method (Gupta & Clarke, 1996; Hung et al., 2008; Landeta, 2006; Landeta & Barrutia, 2011; Linstone & Turoff, 2002; Ziglio, 1996).

As a follow-up to the survey results, I interviewed a sub-group of panelists. Interviews fostered greater understanding from the perspectives of participants (Marshall & Rossman, 2006; Patton, 2002a). I chose participants for the interview portion of the present study in order to explore the responses provided during the Delphi process. In particular, I was interested in any outlier responses and used the interview process to better understand their perceptions. Furthermore, I also interviewed others in the field who did not participate in the Delphi surveys phase but were recommended by others in the study because of their leadership status in the Northeast Florida nonprofit sector or their expertise and insight regarding the subject. Ultimately, the use of interviews provided robust data and fostered deeper understanding regarding the stated research questions regarding the role of and capacity for program evaluation in the Northeast Florida nonprofit sector.

I used content analysis to identify categories, themes, and patterns from an analysis of the qualitative responses from the first exploratory Delphi survey round. I primarily used inductive coding (Marshall & Rossman, 2006; Patton, 2002a) so that the responses of the panelists determined the areas for rating in the subsequent rounds. I analyzed the descriptive quantitative rated responses using frequencies, means, and standard deviations. Data were ranked based on the overall means of categories. I analyzed information from the interviews in the same manner that I analyzed the qualitative responses from the first exploratory Delphi survey round.

In order to have a holistic view of the data, I reported both the qualitative and quantitative data (Tapio, Paloniemi, Varho, & Vinnari, 2011). I used tables to present the quantitative data of frequencies, means, and standard deviations. However, because this is an exploratory qualitative study, I reported in much more detail the qualitative data gleaned from the Delphi surveys and the interviews. Presentation of these data included the results of content analysis for categories, themes, and patterns for both the Delphi surveys and the narrative from the interviews. Additionally, I included information from secondary data (e.g., websites, reports, and annual reports) provided by participants that was relevant to the subject (Marshall & Rossman, 2006).

Assumptions/Delimitations/Limitations

The overall research design and methodology of the present study was limited by certain assumptions. A key assumption of the present study was that the experience of experts in any particular field produces knowledge that has value (B. Johnson & Christensen, 2008). Eisner (1998) described this as “connoisseurship” and argued that all knowledge is based on experience. This is hallmark of qualitative research design and methodology.

Furthermore, an assumption of the present study was that participants chosen for the study responded honestly and that the study participants had specific knowledge and experience regarding program evaluation efforts, and the level of program evaluation capacity, as well as an understanding of the relationships between providers and funders as these relate to the Northeast Florida nonprofit sector. Another assumption of the present study was that the study participants' perceptions have bearing on the reality and/or reflect the reality of the role of and capacity for program evaluation in their individual organizations and the Northeast Florida nonprofit sector as a whole.

Another assumption of the present study was that group affiliation (providers or funders) influences participants' perceptions regarding program evaluation efforts. A closely related assumption was that providers and funders may not communicate regularly regarding all of the varying aspects of the role of and capacity for program evaluation. Additionally, an underlying assumption related to the present study methodology was that honest communication regarding the topic between the two groups may result in negative repercussions for providers (e.g., restricted or loss of access to resources). Consequently, the methodology limited participants from interacting directly. Another assumption of the present study was that program evaluation capacity building has positive outcomes.

Assumptions also shaped the design of the present study. One assumption was that the Delphi methodology would yield useful data to inform the present study. Likewise, I assumed that I, as the researcher, would be able to accurately document the voice of participants. It was also assumed that I would ask the right or best questions for the exploratory Delphi survey round and the interviews. Finally there was an assumption made that I, as the researcher, would be able to bracket and account for my own subjectivity.

Delimitations included restricting the setting for the present study to the five county Northeast Florida area (Baker, Clay, Duval, Nassau, St. Johns counties) and to nonprofit human services and funding agencies. Another delimitation was the restrictive criteria and sampling method used for selecting Delphi panelists and people interviewed for the present study. Another delimitation included the requirement of Delphi panelists to have access to the Internet.

Finally, a delimitation was that the present study occurred during a single point in time. Parameters for the research design and these delimitations are discussed in more detail in Chapter 3 and Chapter 4.

Limitations for the present study included a lack of diversity of participants and construction issues with the Delphi surveys. Another limitation was how participants interpreted instructions in the Delphi survey. These limitations are discussed in more detail in Chapter 4 and Chapter 5.

Definition of Key Terms

For the purpose of the present study, the following operational and commonly accepted definitions were used.

501c3 nonprofit organizations- 501c3 organizations are one of over 25 Internal Revenue Service classifications of tax-exempt status. These are organizations established for religious, charitable, scientific, public safety testing, literary, educational, fostering international amateur sports competitions, and prevention of abuse to animals or children purposes. This category of exempt organizations also includes private foundations. This category also has the largest number of organizations (Internal Revenue Code, 2010; U.S. Department of the Treasury, Internal Revenue Service, 2011; Roeger et al., 2011).

Board of Directors- The governing volunteer board of a nonprofit agency that is legally responsible for the nonprofit organization to which they are affiliated (Axelrod, 2005; Renz, 2010; Worth, 2009).

Chief Executive- The top staff position of a nonprofit organization appointed by the board of directors to carry out the day-to-day operations of the organization. This is typically a compensated position. The two most common titles for this position are

Executive Director or Chief Executive Officer (Herman, 2010; Herman & Heimovics, 2005; Worth, 2009).

Developmental evaluation- Developmental evaluation is a contemporary evaluative purpose posited by Patton (2011) that focuses on "innovation development to guide adaptation to emergent and dynamic realities in complex environments" (p. 1).

Evaluation agenda- An evaluation agenda is a strategic written plan that delineates specifically the purposes, uses, models, methods, and areas of program evaluation for any given nonprofit agency.

Formative evaluation- Formative evaluation includes evaluation efforts for the purpose of improving (B. Johnson & Christensen, 2008) something. In the present study, the focus would be on improving programs of nonprofit organizations.

Funding/philanthropic organizations (funders) – These are organizations or entities that provide funding to nonprofit human services agencies. Typically, funding organizations are nonprofit agencies as well (e.g., foundations, United Way) or government entities at the federal, state, or local levels.

Mission- The mission is the purpose(s) and the "reason for existence" of a nonprofit organization (Knauft, Berger, & Gray, 1991, p. 3).

Nonprofit human services providers (providers)- Nonprofit human services providers are a common designation for a subclass of 501c3 organizations whose missions focus on helping people in need such as the poor, victims of violence, child abuse victims, people in crisis, youth services, services for the elderly, and other such organizations.

Nonprofit organization's programs/service delivery- The services and/or activities provided by a nonprofit organization to its constituents.

<u>Nonprofit sector</u>- The nonprofit sector is the third segment of society, alongside with business and government, that encompasses organizations that typically have tax-exempt status and provide a specific societal benefit/social good to the community. Another key feature is dependence on public and private support for sustainability (Hopkins & Gross, 2010; Worth, 2009).

<u>Organizational accountability-</u> For the present study, the idea that nonprofits have to answer to various stakeholders (e.g., constituents, the public, funders, government, regulators) for various areas such as finances, mission, services, and human resources (Candler & Dumont, 2010; Murray, 2005, 2010).

<u>Organizational effectiveness-</u> A multivariate determination of organizational processes and the degree to which an organization achieves identified organizational goals, such as the mission, in the most efficient manner (Herman & Renz, 1999; Murray, 2005, 2010). For the present study, organizational effectiveness is a sub-component of organizational accountability as the information regarding effectiveness is often provided to stakeholders as a means of accountability.

<u>Perception-</u> For the purpose of the present study, perception is defined as “the way in which something is regarded, understood, or interpreted” (“Perception”, 2013). Belief would be considered a synonym in the present study for perception as Merriam-Webster.com designates the two words as being related (“Belief”, 2013).

<u>Program effectiveness-</u> The degree to which the goals and objectives of a nonprofit organization’s programs/service delivery are realized and the outcome and impact the program has on constituents. This is one part of organizational effectiveness.

<u>Program evaluation capacity</u>- Program evaluation capacity refers to the necessary resources (e.g., funds, knowledge, skills, time) essential to implementing and maintaining quality program evaluation in nonprofit organizations (Carman, 2007; Carman & Fredericks, 2009; Compton et al., 2002; Stockdill, Baizerman, & Compton, 2002).

Quality program evaluation- Quality program evaluation is the purposeful collection of specific information about various aspects of a program to answer questions focused on one or more explicit purposes in a methodical way following standards established by the Joint Committee on Standards for Education Evaluation for multiple purposes of utilization such as validating effectiveness, measuring impact, assessing outcomes, improving services, and informing decision making (Carman, 2007; Compton, Glover-Kudon, Smith, & Avery, 2002; Patton, 1997, 2002a ; Yarbrough, Shulha, Hopson, & Caruthers, 2011).

Summative evaluation- Summative evaluation includes evaluation efforts for the purpose of "determining the overall effectiveness and usefulness" (B. Johnson & Christensen, 2008, p. 11) of something. In the present study, the focus for summative evaluations would be on programs of nonprofit organizations.

Organization of the Study

I organized the present study into five chapters. The first chapter introduces the study and further details the purpose of the study, the significance of the research, research questions, summary of method employed, assumptions, and definition of key terms.

The second chapter is a review of the literature. The review of the literature summarizes information on the state of the nonprofit sector and the theories that informed the conceptual framework of the present study. The theories that influenced the study include the significance of mission to the viability of nonprofit organizations, organizational accountability, program evaluation theory and models, program evaluation capacity, and the importance of collaboration between nonprofit service providers and funding organizations.

Chapter 3 features the methodology used in the present study. Additionally, the chapter details the research design inclusive of the treatment of the data, the description of setting, participant selection, data collection, validity and trustworthiness, the impact of the researcher as instrument, ethical considerations, limitations/delimitations, and a summary.

Chapter 4 includes demographic information regarding the individual participants, represented organizations, detailed information on data analysis, and findings from the study presented through the context of each research question. The chapter concludes with a brief synopsis of the overall findings.

Chapter 5 includes a discussion of the results referencing theoretical frameworks from a review of the literature. Additionally, the chapter includes major conclusions, implications for further research, and suggestions for practice particularly for the Northeast Florida nonprofit sector. The chapter concludes with a summary and reference to the purpose of the present study in the context of the conceptual framework.

Chapter 2: Review of the Literature

The review of the literature summarizes information on the state of the nonprofit sector and the theoretical frameworks that informed the conceptual framework of the present study. The theoretical frameworks included the significance of mission to the viability of nonprofit organizations (mission centrality), organizational accountability, program evaluation theory and models, program evaluation capacity, and the importance of collaboration between nonprofit human services providers and funding organizations.

Nonprofit Sector

The nonprofit sector is a significant part of North American culture representing the fastest growing industry in the United States having experienced a 25% increase in the past decade (Urban Institute, n.d). Nationally, the nonprofit sector represents 9% of income to United States wage earners, employs 10% of wage earners, represents over 5% of the gross national product, engages over 25% of the adult population as volunteers, and garners nearly three billion dollars in private donations (National Center for Charitable Statistics [NCCS], n.d.; Roeger et al., 2011; Urban Institute, n.d.).

Although a plethora of divergent nonprofit agencies exist in the United States, a simple definition for the nonprofit sector is that it is comprised of private organizations established to meet a public need (Mendel, 2010). In the United States, a primary characteristic of nonprofit organizations that differentiates the industry from for-profit business is having tax exempt status (Roeger et al., 2011). In section 501(c-f) of Title 26 of the United States Code, the Federal Government designated several different types of organizations in the United States eligible for tax exemption (see Appendix A for complete list). The most widely known and largest of the exempted categories is the 501(c) 3 designation, which includes charitable organizations and private foundations (Internal Revenue Code, 2010; U.S. Department of the Treasury, Internal Revenue Service, 2011, 2112; Roeger et al., 2011). Additionally, organizations may incorporate with state governments as not-for-profit entities instead of, or in addition to, seeking tax exempt status under the IRS (Gronbjerg, Liu, & Pollak, 2009). For the purpose of the present study, my focus is on IRS tax exempt 501c3 organizations specifically targeting human services agencies, private foundations, and other funding entities.

Although the term nonprofit sector is the most familiar term, other labels are also used to describe the industry such as third sector, independent sector, charitable sector, voluntary sector, tax-exempt sector, civil society sector, community-based organizations (CBOs), private government, nongovernmental organizations (NGO's), social sector, or (more narrowly) 501(3) status (Carman, 2007; Gronbjerg et al., 2009; Hall, 2010; Mendel, 2010; Worth, 2009). For the purpose of this review of the literature, I used the more common designation of nonprofits, nonprofit agencies, or nonprofit organizations to describe the industry.

The roots of the nonprofit sector reach back to the genesis of the colonization of the United States at which time Harvard University and religious organizations formed to assist the disenfranchised (Hall, 2010). However, historically U. S. nonprofits were not categorized as such due to the lack of differentiation between public and private entities. Instead, the culture of the time focused on citizens engaging in collective efforts to meet community goals (Hall, 2010; Mendel, 2010). Foreign visitors noted the phenomenon of the early United States citizenry creating organizations to advance societal objectives (Hall, 2010; Worth, 2009). Alexis de Tocqueville noted, "Americans of all ages, all conditions, and all dispositions constantly form associations . . . . Wherever at the head of some new undertaking you see the government in France, or a man of rank in England, in the United States you will be sure to find an association" (as cited in Hall, 2010, p. 9).

The current designation of the nonprofit sector in the United States as a distinguishable segment of society alongside government and businesses took shape in the last quarter of the 20th century (Hall, 2010; Worth, 2009). Of the three, research by Handy et al. (2010) found that nonprofits were perceived to be the more trustworthy sector as compared to the business and government sectors. The researchers surveyed 1,169 university students using a Likert scale specifically asking the respondents' level of agreement to statements regarding the trustworthiness of nonprofits as compared to business and government. Additionally, the researchers noted that respondents were more likely to demonstrate this trust by supporting nonprofits more through donations and volunteer work. Similarly, in 2011, data collected from a survey initiated by the Nonprofit Center of Northeast Florida provided further evidence of this trust as 67% of the respondents agreed that nonprofits were better situated to help the community than government (Flagg & Rankin, 2011). Ways in which citizens engage the services of nonprofit organizations may include using hospitals, recreational activities, human services organizations, the arts, or educational institutions (Littlepage & Oldakowski, 2006).

Nonprofit Governance

Understanding the governance structure of nonprofit organizations is central to appreciating the unique position the nonprofit sector holds in American society. The governance of a nonprofit organization is a complicated dance between the chief executive and the board of directors. The board of directors in partnership with the chief executive are the top leadership for nonprofit organizations (Axelrod, 2005; Worth, 2009).

By design, the board of directors and the chief executive are an example of shared leadership. Pearce and Conger (2003) provided a definition of shared leadership: "A dynamic, interactive influence process among individuals in groups for which the objective is to lead one another to the achievement of group or organizational goals or both" (p. 1). Further defining characteristics of shared leadership focused on the interactional process among members which produced "mutual learning, greater shared understanding, and eventually, positive action" (Fletcher & Käufer, 2003, p. 23). The organizational structure of a nonprofit organization has shared leadership within the board of directors and with the chief executive. Both the board and the chief executive have unique roles and responsibilities of this interdependent partnership in order to ensure the viability of the organization (Axelrod, 2005; Worth, 2009).

The nonprofit board of directors is typically comprised of volunteers in the community who have an interest in the mission of the organization. The board of directors is not just a group of individuals loosely connected by their involvement with the organization. A nonprofit board acts as a single entity. No particular member can make a decision or speak without consent from the board as a whole. The board elects a member as the chair who serves as the spokesperson for the board of directors (J. Carver, 1997).

The purpose of the board of directors is to safeguard the investment of the community regarding the mission of the nonprofit organization (Axelrod, 2005; J. L. Miller, 2002; Worth, 2009). The legal mandate regarding board of directors' responsibilities includes the duties of care, loyalty, and obedience (Axelrod, 2005; Worth, 2009). The duty of care mandates that board members maintain oversight of the organization's finances and operations through active participation, understanding provided information, making informed decisions, and questioning actions as appropriate. The duty of loyalty means that board members must put the needs of the nonprofit organization above all other considerations. Finally, the duty of obedience requires that board members ensure that the organization obeys all applicable laws and regulations, adheres to the mission, pursues established organizational goals, and practices within the framework of established policies (Axelrod, 2005; Worth, 2009).

Dovetailing with these overarching mandates are the common responsibilities for nonprofit boards. These functions of the board include board development, mission attainment, strategic planning, stewardship of resources, fund-raising efforts, ambassadorship to the greater community, and programmatic accountability (Axelrod, 2005; Bradshaw, Murray, & Wolpin, 1992; W. A. Brown, 2005; Gill, Flynn, & Reissing, 2005; Jackson & Holland, 1998; Nobbie & Brudney, 2003; Renz, 2010). Ultimately, nonprofit boards are the community representatives for public charities (Ostrower, 2007). J. Carver and M. Carver (2009) reported, “The board must first have an adequate vision of its own job. That role is best conceived neither as volunteer-helper nor as watchdog but as trustee-owner” (p. 3).

Although the nonprofit board is the principal leader of an organization, the position of the chief executive does not exist merely at the disposal of the nonprofit board. The board of directors hires or appoints the chief executive to assist with achieving goals and advancing the organizational mission (Axelrod, 2005; Renz, 2010; Sherlock & Nathan, 2007). The chief executive often provides the main source of continuity for the community, board of directors, and other staff members. The chief executive augments stability through mission promotion, setting priorities, and developing key relationships (Worth, 2009). A chief executive hones many skills in order to fulfill the expectations of the position. Sherlock and Nathan (2007) noted several skills such as reflective practice, political acumen, open communication style, and relationship development as core factors for successful leadership.

The chief executive’s most pivotal role relates to assisting the board with responding to its legal responsibilities to the organization (Herman & Heimovics, 2005). Furthermore, the CEO’s effective engagement and development of the board is paramount to success (Worth, 2009). Herman and Heimovics (2005, p. 158) listed several necessary strategies chief executives employed with boards of directors: (a) promoting relationships among board members and the chief executive, (b) being respectful, (c) focusing the board on the future, (d) updating board on key information related to the mission of the organization, (e) providing administrative support and resources for board projects, (f) facilitating achievement of goals and affirmation of success, and (g) planning for executive leadership succession.

Different governance models exist regarding the shared leadership relationship between the chief executive and the nonprofit board (Herman, 2010; Worth, 2009). Models of nonprofit governance elucidate the different strategies used for the leadership of nonprofit organizations. The Carver governance model focuses the attention of the board on developing and monitoring policies. This model creates rigid boundaries between staff and board responsibilities that isolate board members from the everyday reality of the nonprofit (Worth, 2009). In the Carver model of board governance, the chief executive is often responsible for the overall operation of the organization that includes the supervision of other staff, program development, service delivery, fiscal management, fund raising, volunteer management, and public relations (Sherlock & Nathan, 2007; Worth, 2009). However, the board of directors prescribes the degree of freedom that a chief executive has to carry out these duties (J. Carver, 1997; J. Carver & M. Carver, 1997).

Another governance model is the “governance as leadership” model. In this model, the board focuses on fiduciary responsibilities, strategic planning, and creative thinking. The board works in partnership with the chief executive in addressing all three areas in order to be effective leaders. It is essential in the governance as leadership model for the chief executive and the board to participate together in the three areas (Worth, 2009).

An alternative model relevant for board governance is the board-centered leadership model described by Herman and Heimovics (2005). The main task of the chief executive in this approach is to ensure that the board fulfills all legal and public responsibilities. The chief executive makes it a priority to assist the board in developing and maintaining leadership status of the organization.

However, Iecovich and Bar-Mor (2007) found tension for power often existed between the chief executive and the chairperson of the board. More often than not, the chief executive was more dominant in the relationship rather than an equal partner with the chairperson. Similarly, Chait, Ryan, and Taylor (2005) argued that the board of directors and the chief executive often switched leadership responsibilities for the organization, leaving the chief executive with more influence than the board. Yet regardless of the model used for governance, boards that received explicit training regarding roles and responsibilities functioned more effectively with the chief executive and were able to successfully advance the mission of the organization (Nobbie & Brudney, 2003). This is particularly important because the leadership apex as described (board of directors, chief executive) is responsible for ensuring that the nonprofit organization adheres to the mission of the organization, which is central to overall organizational accountability of a nonprofit organization (Knauft et al., 1991).

Accountability to organizational mission, the public, and stakeholders is a primary responsibility of the board of directors of a nonprofit organization and cannot be overstated. Effective nonprofit boards evaluate performance of the overall board, individual board members, executive leadership of the organization, and organizational performance (Axelrod, 2005; J. L. Miller, 2002; Renz, 2010). However, many nonprofit governing bodies do not implement quality evaluation strategies. In spite of the importance of evaluation to the health of nonprofits, Ostrower (2008) concluded,

> Particularly troubling is that almost half of the nonprofit organizations say their boards do not monitor their own performance. Furthermore, more than one-fourth said that their boards do not assess whether the organization is accomplishing its mission, either on an annual or bi-annual basis. (p. 2)

This lack of attention to evaluating board performance and mission attainment is concerning because these are core responsibilities of boards (Worth, 2009) and may impact the success of the organization.

Successful Nonprofits

A review of the literature regarding nonprofits focused on the characteristics of successful nonprofit organizations. Knauft et al. (1991) argued that a high performing nonprofit organization has four essential characteristics: (a) a revered mission for stakeholders to rally around, (b) motivational leaders focused on mission attainment, (c) an active board that worked in shared leadership with the executive officer, and (d) ability to obtain financial support and skilled staff. Letts, Ryan, and Grossman (1999) emphasized the ability to build capacity as an indicator of a high performing organization. Elements of capacity building included investing resources in program development, creating a learning organizational culture, developing staff, engaging key stakeholders, and focusing on expansion. The authors also indicated that it is essential to focus partnerships with funding sources on organizational capacity growth efforts rather than program delivery. Crutchfield and Grant (2008) found that high impact nonprofit organizations were successful advocates for change, promoted shared leadership, engaged stakeholders, created innovative funding opportunities, collaborated successfully with other nonprofit organizations, and were flexible to meet ever-changing external forces (p. 6). Other important characteristics of high performing nonprofit organizations included the importance of investment in technology, use of a strategic plan, data-driven decisions, clearly-defined goals, and use of an effective system of checks and balances (Brinckerhoff, 2000; Light, 2002). The commonality from the review of the literature is that successful nonprofits have strong leadership, are mission-driven, develop strong coalitions with other interested groups, successfully engage stakeholder financial support, and are able to garner other needed resources.

Yet these referenced descriptions of successful nonprofit organizations provided little information on the actual accountability strategies the organizations used to ensure effective service delivery aligned with mission attainment. Indeed, the authors placed scant emphasis on how to institute accountability strategies at all levels of the organization. Although the authors emphasized the need for program evaluation (e.g., program effectiveness, measuring change, bench marking, and performance standards), studying nonprofit organizations without detailing how the organizations measure program effectiveness, outcomes, impact on society, and improvement strategies may be shortsighted.

Nonprofit Accountability

At the time of the present study, nonprofit organizations exist in a climate of extreme accountability. Due to the intense focus on accountability by various stakeholders (e.g., funding organizations, government, and regulators), accountability theory informs the present study and is so crucial to nonprofit operations that experts labeled it the “accountability movement” (Carman, 2007, 2010; Murray, 2005, 2010). Cutt and Murray (2000) defined accountability as

> The obligation to render an account for a responsibility which has been conferred. This definition presumes the existence of at least two parties, one who allocates responsibility and one who accepts it with an undertaking to report on, and account for, the manner in which it has been discharged. (p. 1)

For nonprofit agencies, promotion of organizational mission is a primary component of accountability (Brinckerhoff, 2000; Candler & Dumont, 2010; R. S. Kaplan, 2001). Numerous stakeholders require nonprofits to provide evidence of performance as it relates to mission achievement (Murray, 2005, 2010). However, Murray (2005) noted that the organizational accountability efforts are fraught with issues because of subjective political expectations and tensions.

At the forefront are stakeholders who fund nonprofit organizations such as private charitable foundations, United Way, and all levels of government (local, state, federal). Other stakeholders include regulating agencies, clients, the public, elected officials, volunteers, boards of directors, and nonprofit agency staff (Candler & Dumont, 2010; Carman & Fredericks, 2009; R. S. Kaplan, 2001; Sherlock & Nathan, 2007). Although stakeholders agree about the need for organizational accountability, the criteria for determination are divergent (Forbes, 1998; Herman & Renz, 1999; Newcomer, 1997; Sowa, Selden, & Sandfort, 2004). These often vague factors of accountability make it difficult for nonprofit organizations. Koppel (2005) described this phenomenon as Multiple Accountabilities Disorder (MAD):

> The contention is that the organization suffering from MAD oscillates between behaviors that are consistent with conflicting notions of accountability. The organization will sometimes emphasize the directives of principals, while at other times try to focus on

> customers. In the long run, overseers and constituents are displeased and the organization struggles. (p.95)

Traditionally, accountability for nonprofit organizations concentrated on fiscal measures and business practices (Plantz, Greenway, & Hendricks, 1997). Organizations such as GuideStar, Charity Navigator, the Better Business Bureau, and the Internal Revenue Service use financial information and ratios as a gauge for organizational accountability standards (Coe, 2007; Eckerd & Moulton, 2011; Murray, 2005, 2010; Worth 2009). In the wake of scandals involving public companies and nonprofit organizations, the federal government also instituted tighter financial and governance regulations through the Sarbanes-Oxley Act (J. Carver & M. Carver, 2009; Mendel, 2010; Ostrower, 2007). The government also increased scrutiny of nonprofit organizations' finances through conducting hearings, emphasis on enforcement of current laws, and recommendations for sound governance for nonprofits (Worth, 2009).

Although financial soundness is important, the scope of nonprofit organizational accountability is much more extensive. Herman and Renz (1999) argued that nonprofit organizational accountability indicators "will never be reducible to a single measure" (p. 110) and did not place a lot of stock in program evaluation outcomes as metrics of effectiveness. Yet Worth (2009) linked measures of effectiveness directly to organizational accountability.

Candler and Dumont (2010) noted that measuring organizational performance of a nonprofit agency was complex and difficult. Furthermore, limited understanding, lack of clarity between accountability and effectiveness frameworks, disagreement on applicable measures, and a diversity of definitions of effective performance, and resistance from organizational leaders handicapped meaningful accountability strategies (Baruch & Ramalho, 2006).

Organizational Effectiveness Strategies

Several theoretical frameworks exist that use multi-faceted approaches to assess the effectiveness of different organizational elements. Organizational effectiveness metrics provide information for overall organizational accountability. Worth (2009) highlighted several metrics of organizational effectiveness: financial ratios, using peer benchmarks, mission attainment, program outcomes, and social value. Sowa et al. (2004) suggested, "Given the complexity of the topic, organizational effectiveness should be conceived of and modeled as a multilevel, multidimensional, and structurally integrated concept" (p. 724).

Quinn and Rohrbaugh (1981) introduced the Competing Values Framework (CVF) based on information from several organizational effectiveness theorists. The authors concluded that organizational effectiveness was on a continuum among elements in three broad categories that "compete" for scarce organizational resources. The first category entails the degree to which an organization focuses attention on the good of the people versus the good of the organization. The second category centers on organizational design or structure and the degree to which it is rigid or flexible. The last category of the CVF is the emphasis an organization places on process versus outcomes. However, Herman and Renz (1999) argued that using the CVF is complex.

The dashboard developed by Paton (2003) is another multivariate theoretical framework used for organizational effectiveness. The author prescribed areas of importance for review: (a) current results to include focus on achievements, fiscal information, and marketing information; (b) underlying performance that emphasizes the cost effectiveness of organizational outcomes; (c) risk management at all levels of the organization; (d) an annual assessment of organizational assets to include tangible, human resources, and reputational assets; and (e) keeping the governance body appraised of projects directly influenced by the organization's leadership team

(Paton, 2003, p. 142). Although multidimensional in scope, Paton geared the dashboard towards internal effectiveness strategies but did not emphasize the roles of external stakeholders.

Frameworks for organizational effectiveness provide different strategies to determine effectiveness. Furthermore, determination of organizational effectiveness is a subcomponent of organizational accountability. Much of the information regarding organizational effectiveness can be used on a broader scale to address organizational accountability.

Organizational Accountability Frameworks

Organizational accountability frameworks include a broader focus on to “who” and for “what” nonprofit organizations are responsible (Candler & Dumont, 2010; Ebrahim, 2010; R. S. Kaplan, 2001). Candler and Dumont (2010) posited the non-profit accountability framework focused on accountability to various stakeholders in different areas. Their non-profit accountability framework theory informs the present study through the emphasis on multiple areas of accountability with multiple stakeholders. The authors were very thorough in listing the numerous stakeholders that included members, clients, constituents, donors, government, general public, media, NGO staff, and partners/allies. Areas of accountability in this model included financial resources, volunteers, reputational capital, goods and services, social capital, policy impact, law, formal mission, and ethics (p. 263). The strength of the nonprofit accountability framework is the breadth of areas of accountability cross-referenced with the span of stakeholders.

The balanced scorecard is another theoretical framework for organizational accountability. The balanced scorecard also features elements for determining organizational effectiveness. The balanced scorecard adopts an accommodating multi-dimensional approach centered on accountability to multiple stakeholders such as shareholders, customers, and employees. Furthermore, the balanced scorecard emphasizes learning, adaptation, and change as key strategies (Ahmed, Ahmed, Nawaz, Dost, & Khan, 2011; Chen & Jones, 2009; R. S. Kaplan & Norton, 1992).

R. S. Kaplan and Norton (1992) conceptualized the balanced scorecard framework originally for the business sector. Simplistic in design, yet comprehensive in application, the authors of the balanced scorecard presented the framework as a new way to assess organizations beyond a one-dimensional financial perspective. The components of the balanced scorecard include an analysis of (a) financial perspective that focuses on shareholders and the bottom line; (b) customer perspective which takes into account an organization's standing with consumers; (c) internal business perspective that examines business systems, processes, and products; and (d) innovation and learning perspective that emphasizes product improvement, employee satisfaction, staff development, and valuing of personnel. The authors noted the novelty of their approach:

> The scorecard puts strategy and vision, not control, at the center. It establishes goals but assumes that people will adopt whatever behaviors and take whatever actions are necessary to arrive at those goals. The measurements are designed to pull people towards the overall vision. (R. S. Kaplan & Norton, 1992, p. 79)

As a combined framework for addressing organizational accountability and effectiveness strategies, nonprofit agencies widely adopted and adapted the balanced scorecard approach. The flexibility of use of the balanced scoreboard is apparent in the combined framework for organizational accountability and organizational effectiveness measures. Furthermore, nonprofits often used the balanced scorecard framework (Carman, 2007; Chen & Jones, 2009; Kriemadis, Kotsovos, & Alexopoulos, 2009; Murray, 2005, 2010). R. S. Kaplan (2001) adapted the balanced scorecard for nonprofits by shifting the emphasis to organizational mission rather than the financial interest of shareholders.

The balanced scorecard adapted for nonprofit organizations is a particular theoretical framework of interest for organizational accountability because the framework has been easily adaptable to nonprofits by the realignment of mission as central over revenue. In the nonprofit model of the balanced scorecard, the mission of the organization is central. This is particularly salient because, as noted previously, accountability to organizational mission is the paramount concern for nonprofit organizations (Brinckerhoff, 2000). In order to assure mission integrity, accountability for the nonprofit must focus on meeting outcomes that align with the mission. Charity Navigator introduced this concept through their newest rating component that focuses on how nonprofits report results. The first rating element is “alignment of mission, solicitations, and resources” (Charity Navigator, n.d.).

Additionally, another key component of the nonprofit balanced scorecard is the focus on assessing the effectiveness and impact of organizational activities/services on program participants. These essential and intertwined points of organizational accountability cannot be determined without information produced from a quality program evaluation. The balanced scorecard provides a structure for program evaluation through inclusion of measures, objectives, targets, and initiatives for all key areas (R. S. Kaplan & Norton, 1992).

Accountability and Program Evaluation

Differences exist among organizational accountability, organizational effectiveness, and program evaluation. Organizational accountability is multi-dimensional and a broad concept regarding nonprofits answering to various stakeholders for many different factors regarding the organization as a whole. Organizational effectiveness focuses on demonstrating the health and viability of the total organization through various metrics. Program evaluation is a sub-component of organizational accountability and effectiveness (Murray, 2005, 2010). Program evaluation efforts provide information on overall organizational effectiveness and accountability. Carman (2007) found that nonprofit leaders invested time, finances, and organizational resources in areas of organizational accountability such as audits, regulatory body reviews, licensing requirements, and compliance with accreditation standards. However, they were ignorant regarding the facets of authentic program evaluation. Carman noted:

> These data suggest that although community-based organizations are indeed busy doing activities that we might associate with being more accountable or practicing good management, this does not mean they are conducting evaluation—meaning specific activities that are intended to inform program managers, funders, and evaluators about the results (and processes) of their programs, which would then be used to help improve service delivery. (Carman, 2007, p. 65)

Accordingly, information gleaned from oversight and compliance measures of accountability did not translate to measurement of impact, program effectiveness, or mission obtainment (Carman, 2007). Consequently, program evaluation theory informs the present study, as leaders of nonprofit organizations must understand the diverse roles of program evaluation and the many program evaluation approaches. This knowledge is an essential building block as program evaluation efforts are a vital component of overall organizational accountability.

Program Evaluation

The recorded history of evaluation dates back thousands of years. Ancient Chinese documents recorded evaluations of personnel (Mark, Henry, & Julnes, 2000). The biblical account of creation documented God as an evaluator of creation, “God saw all that he had made, and it was very good" (Genesis 1:31, New International Version; Patton, 1997). Patton (2002a) also detailed an account of a systematic evaluation program documented in the book of Daniel in the Bible. Indeed, evaluation seems an inherent aspect of the human experience (Patton, 1997). The genesis of program evaluation as a profession in the United States dates back to the 1960s when various social service initiatives funded by the federal government were required to account for the effectiveness of their programs (Chelimsky, 1997; Chelimsky & Shadish, 1997; Duignan, 2003; Febey & Coyne, 2007; Madaus & Stufflebeam, 2000; Mark et al., 2000). Influenced by education, human resources, business, the research community, and growing demand for accountability, program evaluation developed into a distinct discipline (Mark, et al., 2000; Patton, 1997, 2002a; Scriven, 1981; W. K. Kellogg Foundation, 2004a). Furthermore, Urban and Trochim (2009) posited that evaluation is a natural bridge between the research and practice communities, making evaluation pivotal to both.

Scriven (1981) defined evaluation as "the process of determining the merit, worth and value of something; or the product of that process" (p. 53). A broadened definition offered by Patton (2010) is "the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming" (p. 13). In addition to these two definitions for evaluation is a range of thought regarding the objectivity or subjectivity of evaluation. On one hand, Scriven prescribed a more positivistic view of evaluation as being objective (Chelimsky, 1997; Dugan, 1996). However, detractors of a value-free operational definition of evaluation argued that evaluation can never be free of subjectivity (Chelimsky, 1997; Stufflebeam, 2001). Evaluation may occur across many disciplines and levels of inquiry such as organizations, programs, projects, human resources venues, systemic analysis, policies, goods, and services (Davidson, 2005; Mark et al., 2000; Scriven, 1981).

Roles of Program Evaluation

The three types of evaluative purposes are formative, summative, and the more recent developmental (Patton, 2011). Formative and summative are the most common and traditional groupings for evaluative purposes (Arnett, 1993; Davidson, 2005; Newcomer, Hatry, & Wholey, 2010; Patton, 1997, 2002a). Scriven (1981) reasoned that formative evaluation centers on improvement while summative evaluation facilitates decision-making. Developmental evaluation is a contemporary evaluative purpose posited by Patton (2011) that focuses on "innovation development to guide adaptation to emergent and dynamic realities in complex environments" (p. 1). Patton differentiated this evaluative purpose from formative and summative evaluations because of the emphasis on information to advance change.

The use of summative evaluation may include descriptions of organizational achievements or shortcomings, prescribing value, funding determination, accountability to funders or governing bodies, appraisal of effectiveness, and measurement of causal relationships among variables (Chelimsky, 1997; Davidson, 2005; Mark et al., 2000; Patton, 1997, 2002a; Scriven, 1981). Developmental evaluation specifically includes facilitating the needed evolutional change of systems, programs, or policies in order to advance the organizational mission (Patton, 2011). Ultimately, the common purposive theme among these three types of evaluation is an assessment of the product, service, or process under scrutiny for the betterment of the beneficiaries of the program under evaluation (Patton, 2002a; Posavac & Carey, 2003).

Explicit benefits of program evaluation are related to the more distinct roles for which program evaluation is used in the nonprofit sector. Eckerd and Moulton (2011) distinguished three main areas of concentration for the role of program evaluation in nonprofits:

> These uses of evaluation can be categorized into three purposes: instrumental, where the results directly change the organization's behavior; conceptual, where results indirectly change the organization through learning; and symbolic, where results are used purely for signaling purposes and no change occurs. (p.101)

More pointedly, one of the roles of program evaluation in the nonprofit sector is assessing the alignment of activities and outcomes with organizational mission (Eckerd & Moulton, 2011; Murray, 2005). This role of program evaluation is consistent with the emphasis placed on organizational mission in the previously noted accountability frameworks (Candler & Dumont, 2010; R. S. Kaplan, 2001). Furthermore, boards of nonprofits are responsible for ensuring mission alignment (Axelrod, 2005; Worth, 2009).

A review of the literature detailed other roles of program evaluation in the nonprofit sector. A most often cited role of program evaluation is to provide information for reports to stakeholders such as funders, board members, regulatory entities, and the general public (Carman & Fredericks, 2009; Cousins et al., 2008). Other reasons for and the benefits of instituting quality program evaluation in nonprofit organizations go beyond the requirements of regulatory and funding bodies (Eckerd & Moulton, 2011; Mark et al., 2000; Stevenson, Florin, Mills, & Andrade, 2002). As mentioned previously, using information to improve programs is another reason for implementing evaluation strategies (Fetterman, 1996; R. L. Miller & Campbell, 2006; Patton, 2002a; Stufflebeam, 2001; 2007). An additional key reason cited for implementing program evaluation includes promoting a learning environment in the organization (Patton, 1999, 2007, 2010; Stevenson et al., 2002; Woodwell, 2005).

Additional roles of program evaluation beneficial to the nonprofit sector include (a) quality assurance assessment, (b) increasing knowledge/learning, (c) outcomes measurement, (d) determining program effectiveness, (e) program innovation, (f) assisting with management decisions, (g) increasing revenue, (h) legitimizing organization, (i) assistance with strategic planning , and (j) assessment of impact (Carman, 2007, 2011; Carman & Fredericks, 2008, 2009; Cousins et al., 2008; Kehrer, 1993). These different roles of program evaluation are not unique to the nonprofit sector. Lee, Altschuld, and Hung (2008) found similar purposes for the role of program evaluation in the educational sector.

Linked to these different purposes for program evaluation in the nonprofit sector are different models or approaches. A variety of approaches exists that focus on the different purposes or needs for an evaluation. Understanding the different approaches is important to developing an evaluation agenda that best meets the needs of any given organization.

Program Evaluation Approaches

Different approaches serve different purposes of program evaluation for the nonprofit sector. Many approaches (models) of program evaluation theory inform the present study for the purpose of demonstrating the plethora of options from which nonprofits may choose to best suit their evaluation purposes and provide data for overall organizational effectiveness and accountability. Stufflebeam (2001) categorized the different approaches into four broad categories: pseudo evaluations, questions and methods oriented, improvement/accountability, and social agenda/advocacy. Others categorize evaluation approaches according to the focus of the evaluation such as need, process, outcome, or efficiency (Posavac & Carey, 2003). Still others categorize evaluations based on the method used such as quantitative, qualitative, or mixed (B. Johnson & Christensen, 2008). I chose to categorize evaluation approaches into five categories adapted from Patton (1997) and Stufflebeam (2001): knowledge-focused, oversight-focused, impact-focused, comprehensive-focused, and participatory-focused. I assigned an approach to one of these overarching categories based on previous cataloging of approaches by experts in the field and relevance to use in nonprofit organizations. Groupings also reflect the implied application or use of the information rather than any similarities of evaluation methods employed.

Knowledge-focused evaluation. The knowledge-focused category as described by Patton (1997) parallels those approaches traditionally used in other disciplines such as the social sciences and anthropology. The primary characteristics are providing information and increasing understanding. Evaluators procure knowledge for knowledge's sake and actions based on the information are not required. Evaluators use knowledge-focused evaluations to understand program theory, conduct assessments, explain preferences or differences, increase understanding, clarify, define purposes, influence thought, and to provide an overview of a subject (Arnett, 1993; Patton, 1997; Stufflebeam, 2001). Included in the knowledge-focused category is traditional research, the naturalistic approach, and program theory approaches.

The research approach employs experimental and quasi-experimental designs popular with social science disciplines (B. Johnson & Christensen, 2008; Mark et al., 2000; Posavac & Carey, 2003; Stufflebeam, 2001). Researchers commonly use the naturalistic approach in anthropology by evaluating phenomena in its natural setting to promote understanding of groups or cultures (Arnett, 1993; B. Johnson & Christensen, 2008; Patton, 2002a; Posavac & Carey, 2003). The program theory approach describes, typically in a linear fashion, how a program will achieve expected results. Evaluators and program planners often use logic models to provide a structure for evaluation efforts in the program theory approach, and many funding organizations require nonprofit organizations to use logic models in grant requests (Patton, 2010; Posavac & Carey, 2003; Stufflebeam, 2001; Urban & Trochim, 2009; W. K. Kellogg Foundation, 2004a, 2004b). The W. K. Kellogg Foundation (2004a) provided a definition:

> A program logic model is a picture of how your program works – the theory and assumptions underlying the program. A program logic model links outcomes (both short- and long-term) with program activities/processes and the theoretical assumptions/principles of the program. This model provides a roadmap of your program, highlighting how it is expected to work, what activities need to come before others, and how desired outcomes are achieved. (W. K. Kellogg Foundation, 2004a, p. 35)

The concept behind use of the logic model is that it will allow practitioners to clearly think out program strategies and outcomes in a logical and systematic manner.

Oversight-focused evaluation. I categorized the next group of approaches as oversight-focused and the main characteristic is the assessment of compliance to prescribed standards, governing bodies, and program funders. Other common characteristics include using the information from the evaluation for quality assurance measures. The information is also used to make administrative and funding decisions. The oversight-focused classification includes audit, applied management, fiscal, consumer/product, and industrial inspection approaches.

The audit evaluation approach is more consistent with program monitoring, assurance of standards, and an assessment of whether the program generates a return on the investments made (Posavac & Carey, 2003; Stufflebeam, 2001). The applied management approach includes evaluations conducted for providing management with feedback in order to make decisions (Arnett, 1993; Stufflebeam, 2001). Areas of concern for management may include budgets, program planning, human resources, reporting, performance, organizational allotment of resources, direction of organizations, and other functions of management (Fayol, 1916/2005; Gulick, 1937/2005). The fiscal evaluation approach concentrates on assessing if the program is producing enough “bang for the buck” through use of cost-benefit or cost-effectiveness analysis (Arnett, 1993; Posavac & Carey, 2003; Stufflebeam, 2001). The consumer/product is a traditional approach (Posavac & Carey, 2003) where an independent evaluator determines the “merit or worth or value of something” (Scriven, 1981, p. 53). Scriven is the most noted contributor of the approach and his approach is often described as being objective and value-free (Dugan, 1996; Posavac & Carey, 2003; Stufflebeam, 2001; Scriven, 1981). Finally, the industrial inspection report is a basic inspection after production and used in factories or other businesses where an item is manufactured (Posavac & Carey, 2003). Although classified by professional evaluators and theorists as evaluation models, the oversight-focused approaches noted are summative in ideology and more characteristic of accountability frameworks (mentioned in the previous section) rather than program evaluation models.

Impact-focused evaluation. I label the next class of evaluations as impact-focused approaches. At the heart of this group are determination of program effectiveness and extent of program outcomes, which are directly related to some the roles of program evaluation in the nonprofit sector. The impact-focus classification is similar to the judgment-oriented evaluation category defined by Patton (1997) as an approach that “requires preordinate, explicit criteria and values that form the basis for judgment” (p. 68). Patton noted that funders and stakeholders outside of the organization are more oriented to this form of evaluation. A distinct difference between this category of approaches and the latter (oversight-focused) is that impact-focused evaluations provide more information on results and the overall efficacy of a program. Those who use oversight-focused evaluation are more concerned with compliance and the bottom line.

The impact-focused group includes student achievement, objectives-based, impact/outcome, goal-free, expert/connoisseurship, clarification hearings, and the Friedman performance accountability model approaches. Student achievement approaches generate from the long history of academic evaluations often using standardized tests (Stufflebeam, 2001). The objective-based approach assesses the degree to which an organization reached stated objectives or declared deliverables (Arnett, 1993; Posavac & Carey, 2003; Stufflebeam, 2001). Impact/outcome evaluation seeks to uncover the overall effect the program has on constituents (Arnett, 1993). Similar to the impact/outcome approach, the goal-free purpose is to determine program impact. However, it differs from the latter as evaluators do not know the intended outcomes and therefore discover all positive and negative outcomes whether intended or not (Arnett, 1993). The expert/connoisseurship approach uses an expert in the field of study as an evaluator. The expert provides feedback and makes overall assessments (Arnett, 1993; Eisner, 1998; Posavac & Carey, 2003; Scriven, 1981; Stufflebeam, 2001).

Mark Friedman (2005) created an evaluation model based on three essential questions: (a) How much did we do? (b) How well did we do it? and (c) Is anyone better off? (p.67). These questions indicate a desire to understand the impact of service delivery on the consumer, which is the reason I listed it in this category. The questions are in a grid with program personnel deciding how to answer and quantify each question. The model is particularly relevant as it is an evaluation approach The United Way of Northeast Florida requires for nonprofit grantees. However, Friedman heavily promoted using survey results from participants as a measure of results, which may skew results or not accurately measure impact. The last approach in this category is the clarification hearing which models a trial. Evaluators argue the pros and cons of the point of focus and conclude worth by which side put forth the best argument (Stufflebeam, 2001).

Comprehensive-focused evaluation. I grouped the fourth class of evaluations as comprehensive-focused approaches. The characteristics of this group are evaluation of multiple areas, emphasis on programmatic processes, and focus on improvement. Approaches in this group include CIPP (Stufflebeam & Wingate, 2005), decision-oriented, and process models of evaluation. Patton (1997) summarized that "what these approaches share is a focus on improvement—making things better—rather than rendering summative judgment… gathering varieties of data about strengths and weaknesses … and each can be used to inform an ongoing cycle of reflection and innovation" (p. 68).

CIPP is a comprehensive evaluation approach developed by Daniel Stufflebeam and combines focus on process as well as impact (Arnett, 1993; Coryn, Schroter, & Hanssen, 2009; Stufflebeam, 2000; Stufflebeam, 2001). Defined by Stufflebeam (2007):

> The CIPP Evaluation Model is a comprehensive framework for guiding evaluations of programs, projects, personnel, products, institutions, and systems… Corresponding to the letters in the acronym CIPP, this model's core parts are context, input, process, and product evaluation. In general, these four parts of an evaluation respectively ask, What needs to be done? How should it be done? Is it being done? Did it succeed? (p. 1)

Closely related to CIPP is the decision-oriented approach which Stufflebeam (2001) coined as "decision/accountability-oriented" (p. 42). Posavac and Carey (2003) succinctly labeled this approach as "improvement-focused" (p. 28). The main point is the combination of summative and formative evaluative purposes to detail the merit of the program and provide information and feedback for program improvement.

Akin to the decision-oriented approach are process evaluations. As labeled, the focus of process evaluations is understanding the processes used. Evaluators often use process evaluations to inform program improvement as well as complement an impact/outcome evaluation (Arnett, 1993). I grouped them in the comprehensive evaluation category because the emphasis is on evaluation of the process and program improvement rather than outcomes.

Participatory-focused evaluation. I adapted Stufflebeam's (2001) advocacy-focused classification of evaluation approaches in order to create the last group as participatory-focused. I grouped client-centered/responsive, constructivist, deliberative democratic, utilization-focused, and empowerment evaluation approaches in the participatory-focused cluster. Common characteristics of participatory-focused approaches are meaningful participation of key stakeholders and inclusiveness of a broad spectrum of disenfranchised groups, promotion of equality, advancement of subjective knowledge, betterment of society, and addressing social justice issues (Chelimsky, 1997; R. L. Miller & Campbell, 2006; Tang et al., 2002).

The client-centered/responsive evaluation approach "levels the playing field" among stakeholders, promotes program improvement, and provides tools for constituents to evaluate their own programs (Stufflebeam, 2001). As an overarching philosophy, proponents of constructivist evaluations maintain that information produced from evaluations is always subjective, does not necessarily reflect conclusions of all stakeholders, and is not definitive. In client-centered approaches, the evaluator leads the evaluation but works closely with stakeholders (Stufflebeam, 2001). Similarly, deliberative democratic evaluators focus on including stakeholders. Ultimately, however, the evaluator determines the final claims and outcome of the evaluation.

Patton (2002b) conceptualized utilization-focused evaluations as being “judged by their utility and actual use; therefore, evaluators should facilitate the evaluation process and design any evaluation with careful consideration of how everything that is done, from beginning to end, will affect use” (p.1). The main point of the approach is that the customer use the information gleaned from the evaluation to inform and improve programs. Patton (1997) argued that evaluations not used by program stakeholders serve minimal purpose. Although noted as an approach, Patton (1997) contended that “utilization-focused evaluation does not advocate any particular evaluation content, model, method, theory, or even use. Rather, it is a process for helping primary intended users select the most appropriate content, model, methods, theory, and uses for their particular situation” (p. 1). I grouped this approach in the participatory category because the emphasis Patton places on involving the client in all stages of the evaluation in order to produce an evaluation used by the client.

Empowerment evaluation as a model is unique from the previous three approaches as it views the evaluator’s role as that of a technical assistant and as an “agent of social change” (R. L. Miller & Campbell, 2006, p. 297) rather than a leader of the evaluation. The degree of self-determination of stakeholders is of prime value to the empowerment-focused evaluator (Stufflebeam, 2001). Consequently, Fetterman (1996) defined empowerment evaluation as “the use of evaluation concepts, techniques, and findings to foster improvement and self-determination” (p. 4). A key benefit of empowerment evaluation as a model is capacity building for ongoing evaluation (Mayer, 1996). However, Patton and Scriven argued that empowerment evaluation is not distinct enough to label it as a specific approach to evaluation (R. L. Miller & Campbell, 2006). Refer to Table 1 for a summary of program evaluation approaches.

Program Evaluation Standards

Understanding the accepted standards under which these models operate is also important for the nonprofit sector. Implementation of these models using industry standards of excellence facilitates quality program evaluations. The Joint Committee on Standards for Educational Evaluation (Yarbrough et al., 2011) listed five major standards areas: (a) proper utilization of

Table 1

Summary of Program Evaluation Categories and Approaches

Categories	Descriptors	Approaches
Knowledge-Focused	• Increase understanding • To inform • Describe and assess	• Traditional research • Naturalistic approach • Program theory
Oversight-focused	• Compliance monitoring • Quality assurance • Administrative decisions	• Audit • Applied management • Fiscal • Consumer/product • Industrial inspection
Impact-focused	• Effectiveness • Results/outcomes • Value	• Student achievement • Objectives-based • Impact/outcome • Goal-free • Expert/connoisseurship • Clarification hearings • Friedman performance accountability model
Comprehensive-focused	• Multi-faceted • Improvement • Processes and activities	• CIPP • Decision-oriented • Process
Participatory-focused	• Stakeholder inclusiveness • Active engagement • Subjective knowledge	• Client-centered/responsive • Constructivist, • Deliberative democratic • Utilization-focused • Empowerment evaluation

evaluations; (b) using evaluations that are feasible and efficient; (c) that evaluations conform to legal, ethic, and moral mandates with highest respect to participants; (d) information is transparent and accurate; and (e) provides levels of accountability. Within these broad categories are 30 subcategories providing more specific information. Consequently, distinguishing the different models and level of quality based on established standards provides the context for understanding the role of and capacity for program evaluation in the nonprofit sector.

In summary, knowing the different evaluative purposes, various program evaluation approaches, and the beneficial roles of program evaluation is critically important to understanding the many options available to nonprofits. In fact, the field of evaluation with the numerous evaluation purposes, models, and standards can be overwhelming for untrained nonprofit staff. Furthermore, describing the depth and breadth of quality program evaluation choices provides a context for understanding the feasibility and impediments faced by nonprofit organizations attempting to determine the role of and capacity for program evaluation in the nonprofit sector.

Nonprofit Organizations and Program Evaluation

The push for measuring the outcomes and impact of programmatic activities began in the mid 1990s (Hendricks et al., 2008; Herman & Renz, 1999; Newcomer, 1997). The United Way of America was one of the first major funding organizations that required some type of program evaluation (Carman & Fredericks, 2009; Murray, 2005, 2010). Another influence during the same period was the Government Performance and Results Act of 1993, which introduced public agencies to the use of logic models for evaluative purposes (Behrens & Kelly, 2008). The trend for data regarding outcome measures, effectiveness, and program impact continued as more private donors, government entities, foundations, and the general public required information on the efficacy of programs delivered by nonprofit organizations (Behrens & Kelly, 2008; R. Brown & Reed, 2002; Carman, 2007; Carman & Fredericks, 2009; Cousins et al., 2008; Garcia-Iriarte, Suarez-Balcazar, Taylor-Ritzler, & Luna, 2011; Kegeles et al., 2005; Sowa et al., 2004; Urban & Trochim, 2009).

The lack of evaluation capacity in nonprofit organizations is the most glaring barrier to instituting a quality program evaluation agenda. Several recent publications report that the most noted elements of inadequate program evaluation capacity in nonprofit organizations were the lack of funding, a hostile organizational culture regarding program evaluation, and human resources deficiencies (Carman, 2010; Carman & Fredericks, 2009; Garcia-Iriarte et al., 2011; Kegeles et al., 2005).

Funding

The lack of adequate funding for implementing quality program evaluation strategies in nonprofit organizations is a perennial issue. Funders such as government, private foundations, and the United Way require evaluation efforts (Hendricks et al., 2008) but often lack understanding of the cost of evaluation or rarely support nonprofits in building the capacity to implement quality program evaluation (Carman, 2010; Duignan, 2003; Kehrer, 1993). Mandates by funders have often created ineffective evaluation strategies and may compromise collaboration between funders and nonprofit organizations (Carman & Fredericks, 2009). Additionally, funders often require evaluation as a monitoring tool rather than an opportunity to promote organizational learning and program improvement, which often deflected nonprofit organizations from using program evaluation for quality assurance purposes (Carman 2007, 2011). Ebrahim (2010) concluded

Evaluations or performance assessments that reward success while punishing failure (for example, through revocation of funds or additional conditions on funding) seem unlikely to engender learning since they encourage nonprofits to exaggerate successes while discouraging them from revealing and closely scrutinizing their mistakes. (p. 113)

The lack of adequate funding has also compromised organizations' ability to procure needed technology such as up-to-date hardware and software programs designed for program evaluation (Carman & Fredericks, 2009; Hendricks et al., 2008; Stevenson et al., 2002). Consequently, Carman (2007) indicated that most nonprofit organizations have made funding program evaluation efforts a low priority. Yet organizations spent funds on audit and monitoring activities. Furthermore, Kegeles et al. (2005) reported resistance and resentment among nonprofit organizations for using limited resources to fund program evaluation efforts.

Organizational Culture

The elements of organizational culture theory relevant to the present study include the dominant attitude regarding program evaluation in nonprofit organizations and the degree to which organizational culture influences successful implementation of program evaluation strategies. Schein (1993/2005) defined the broader organizational culture as "a pattern of shared basic assumptions that the group learned as it solved its problems… taught to new members as the correct way to perceive, think, and feel in relation to those problems" (pp. 364-365). Wu (2008) characterized organizational culture "as an evolutionary process where people within an organization learn from making repeated choices." (p. 2538).

The way an organization responds to program evaluation efforts is a narrower and more specific component of the overall organizational culture. Kegeles et al. (2005) found that a negative organizational culture around program evaluation may serve as a key barrier to implementation. As identified in a review of the literature, elements of organizational culture that hinder quality program evaluation efforts included (a) a low level of cooperation from volunteers, staff and other key stakeholders; (b) a high level of resistance to change from key stakeholders; (c) a low level of perceived need for evaluation by stakeholders; (d) a high level of adversarial office politics/power struggles among various stakeholders; and (e) divergent evaluative purposes among collaborators (Behrens & Kelly, 2008; Bozzo, 2002; Hendricks et al., 2008; Stevenson et al., 2002). Alaimo (2008) also reported that positive (or negative) organizational attitudes towards evaluation started with the chief executive and that support at the executive level was necessary for successful implementation of a quality program evaluation agenda.

The time an organization allocates for program evaluation activities may also be an indication of the organizational culture regarding program evaluation efforts. For example, as noted earlier, nonprofit organizations frequently allocate more time for monitoring, ensuring successful audits, and compliance with regulatory standards than program evaluation efforts (Carman, 2007). Studies have also noted that the lack of available time for program evaluation efforts is often a barrier (Carman & Fredericks, 2009; Hendricks et al., 2008; Kegeles et al., 2005). Furthermore, nonprofits typically face high staff turnover, and training new staff regarding program evaluation efforts takes time (Hendricks et al., 2008; Kegeles et al., 2005). Likewise, T. I. Miller, Kobayashi, and Noble, (2006) identified the use of staff time to conduct evaluation efforts as a high cost factor to an organization.

Organizational design is clearly related to organizational culture, and a poorly designed organization is likely to struggle with program evaluation capacity and implementation. Organizational design includes internal structures, size of the organization, processes, stability, and the types of organizational hierarchy not supportive of program evaluation (Carman &

Fredericks, 2009; Kegeles et al., 2005; Stevenson et al., 2002; Stockdill et al., 2002). A study by Carman and Fredericks (2009) found that larger organizations had more difficulty implementing and sustaining program evaluation. Kegeles et al. (2005) reported that smaller nonprofit organizations were more resistant to program evaluation efforts due to limited resources and negative attitudes regarding program evaluation activities. The common feature between the larger and smaller nonprofits from both of these studies was overall organizational capacity struggles such as staff issues and lack of needed resources. These systemic capacity problems were also were indicative of program evaluation capacity deficits.

Human Resources

Researchers have frequently recognized that human resources are clearly linked to an organization's capacity for program evaluation. Specific human resource issues include high staff turn-over, limited number of staff to perform evaluation activities, and the lack of skills and knowledge regarding program evaluation (Carman & Fredericks, 2009; Eckerd & Moulton, 2011; T. I. Miller et al., 2006; Stevenson et al., 2002). Studies indicated that nonprofit managers and frontline staff had very limited knowledge of program methods or designs. Employees did not understand program evaluation design and the link to developing goals, objectives, or outcomes (Kegeles et al., 2005). They did not know the different evaluation approaches available. They were unclear about industry standards of excellence. Furthermore, staff often lack sufficient data, may not know about different measurement tools available, and are often unsure about how to analyze the data or how to use the data once compiled (Hendricks et al., 2008; Kegeles et al., 2005; Stevenson et al., 2002). In fact, nonprofit leaders erroneously concluded that auditing, monitoring, or compliance activities were program evaluation strategies (Carman, 2007; Carman & Fredericks, 2009).

Another key factor that has been found to hamper capacity of people is the limited use of professional evaluators by nonprofit organizations (Carman, 2007; Kegeles et al., 2005). However, there is considerable debate regarding use of external evaluators. Some program evaluation experts have promoted the idea that program evaluation rests best in the hands of the professional evaluators (R. M. Johnson, 1993; Newcomer, 2001). Others have argued that evaluation experts, funders, and program staff need partnerships with each other to implement quality program evaluation (Arnold, 2006; Hendricks et al., 2008; Kegeles et al., 2005; Mayer, 1996). Still others have claimed that nonprofit organizations should effectively train internal staff to implement evaluation strategies (Duignan, 2003; W. K. Kellogg Foundation, 2004a). However, external stakeholders have often viewed self-evaluations by nonprofit organizations as somewhat suspect (T. I. Miller et al., 2006).

In summary, program evaluation is essential to the success of nonprofit organizations. However, a review of the literature suggests that the level of program evaluation capacity has a direct bearing on program evaluation efforts. The literature also reflects essential program evaluation capacity components and how the development of capacity is vital to program evaluation efforts.

Program Evaluation Capacity

The theoretical literature on program evaluation capacity building served to inform the present study. The lack of program evaluation capacity is a barrier for implementing quality program evaluation in nonprofit organizations. Adequate program evaluation capacity is essential to designing, implementing, and sustaining evaluation agendas in nonprofit organizations. Carman (2007) concluded that "nonprofit organizations need to start investing in their own evaluation capacity, much in the same way that they have in other critical management functions, such as financial management, accounting, fundraising, donor development, and volunteer management" (p. 73). Stockdill et al. (2002) succinctly defined evaluation capacity building (ECB) as "the intentional work to continuously create and sustain overall organizational processes that make quality evaluation and its uses routine" (p. 14). As simple as the definition appears, evaluation capacity building is multi-layered (Compton et al., 2002), requires considerable organizational resources, and needs time to be established (Poole, Nelson, Carnahan, Chepenik, & Tubiak, 2000).

Evaluation Skills Development

Development of evaluation skills among the various stakeholders in nonprofit organizations is paramount to building evaluation capacity. First and foremost, inclusion of all essential stakeholders in assessing, developing, implementing, and maintaining evaluation capacity is critical (Yarbrough et al., 2011). Stakeholders include program participants, the public, staff, volunteers, board members, funders, regulating entities, and others that may have a vested interest in the organization (Bryson & Patton, 2010; Cousins et al., 2008; Milstein, Chapel, Wetterhall, & Cotton, 2002; Tang et al., 2002). Inclusion of multiple stakeholders is particularly important when designing an evaluation agenda as stakeholders vary in the criteria they use in determination of success (Coryn et al., 2009). After identification of key stakeholders, it is vital to designate selected stakeholders (typically staff) to become evaluation specialists (Arnold, 2006; King, 2002). It is also important that the stakeholders chosen as evaluation specialists be receptive to learning or strong in advocating for a sustained evaluation agenda (Garcia-Iriarte et al., 2011; Milstein et al., 2002). Increasing the skill levels of specialists and other crucial stakeholders is also necessary to build program evaluation capacity (Behrens & Kelly, 2008; Duignan, 2003). A review of the literature emphasized fundamental areas for increased knowledge: (a) understanding the importance of evaluation (Kegeles et al., 2005); (b) understanding the differences of evaluation elements such as outputs, objectives, and outcomes; (c) using different evaluative approaches as needed (Duignan, 2003; Milstein et al., 2002); and (d) understanding which type of approach to use for the given context (Behrens & Kelly, 2008). Furthermore, Arnold (2006) posited that foundational knowledge and use of logic models in organizations are indispensible to quality program evaluation.

Logic models are often connected with program theory, theory of change, and theory of action (Funnel & Rogers, 2011; Patton, 1997, 2002a). Funnel and Rogers (2011) defined program theory as "an explicit theory or model of how an intervention contributes to a set of specific outcomes through a series of intermediate results" (p. 31). The authors (Funnel & Rogers, 2011) also defined the theory of change as "the central processes or drivers by which change comes about for individuals, groups, or communities" (p. 34). Additionally, Funnel and Rogers (2011) defined the theory of action as "how programs or other interventions are constructed to activate their theory of change" (p. 31). According to Patton (2002a), the difference between the theory of change and the theory of action "is that the theory of change is more research based and scholarly in orientation, whereas a theory of action is practitioner derived and practice based" (p. 163). In short, the theory of action details how a program produces desired results. Patton distinguished the theories of change and action from logic models in that "logic models are descriptive. Theory of change and theory of action models are explanatory and predictive" (Patton, 2002a, p. 163). However, Patton also used the term "logic model" as a synonym for theory of action (Patton, 2002a; 2010). Essentially, logic models are graphic and descriptive representations of the theory of action (Patton, 2002a; Taylor-Powell & Henert, 2008). Logic models also become a component of evaluation (Patton, 1997, 2002a;

Thomas, 2010; W. K. Kellogg Foundation, 2004b).

Logic models map out the steps a program uses to accomplish change. In essence, "a logic model brings program concepts and dreams to life" (W. K. Kellogg Foundation, 2004b, p. 3). Experts contend that understanding and appropriately using logic models to detail program service activities are fundamental evaluation skills for nonprofit staff (Arnold, 2006; Garcia-Iriarte, et al., 2011; Poole et al., 2000). However, Herman and Renz (1999) suggested that placing attention on logic models and outcomes is risky as it overly emphasizes limited measures of effectiveness as an indication of organizational accountability.

Yet funders often require logic models in grant applications (Behrens & Kelly, 2008; Compton et al., 2002; Hendricks et al., 2008; W. K. Kellogg, 2004a, 2004b). Many variations of logic models exist (Taylor-Powell & Henert, 2008). However, the W.K**.** Kellogg Foundation (2004b, p. 3) summarized the logic model into five simple steps:

1. A listing of the different resources/inputs needed for program implementation.
2. A listing of the types of expected activities.
3. A listing of the outputs, services or product resulting from activities.
4. A listing of the intended outcomes/results achieved because of the activities.
5. A listing of overall desired impact or change because of the outcomes produced by the activities.

Many uses exist for logic models such as program planning, program implementation, formative and summative evaluation purposes, communication, establishing clear outcomes, and a tool for strategic planning (S. A. Kaplan & Garrett, 2005; McLaughlin & Jordan, 1999; Taylor-Powell & Henert, 2008; Urban & Trochim, 2009; W.K. Kellogg Foundation, 2004b).

Creating an Evaluation Culture

Kegeles et al. (2005) indicated that one of the primary areas related to program evaluation capacity is organizational culture. Cultural shifts in attitudes towards instituting quality evaluation strategies need to occur at all levels of an organization (Kegeles et al., 2005; Stevenson et al., 2002). Furthermore, supportive leadership from the head of the organization and the management team are needed to move the organizational culture towards evaluation (Kegeles et al., 2005; Milstein et al., 2002).

A review of the literature reported specific useful strategies that positively influence organizational culture regarding evaluation measures. Focus groups with employees have been shown to be helpful in understanding concerns and issues (Milstein et al., 2002). Communication regarding the relevance of evaluation to everyday work and program goals is also essential (King, 2002; Milstein et al., 2002). Increasing staff understanding of evaluation and its related benefits is also effective for building program evaluation capacity (Duignan, 2003; Kegeles et al., 2005). Increased training, mentoring, and creation of manuals regarding evaluation for all staff can also positively influence organizational culture (Duignan, 2003; Garcia-Iriarte et al., 2011; Hendricks et al., 2008; Milstein et al., 2002; Poole et al., 2000). Moreover, specifically identifying and encouraging a person(s) in an organization who supports evaluation efforts as a means to influence the attitudes of others may help establish and maintain a positive culture towards evaluation (Garcia-Iriarte et al., 2011). These strategies have been shown not only to influence organizational culture towards instituting evaluation but also to promote a reflective practice often found in learning organizations (Baizerman, Compton, & Stockdill, 2002).

Cook and Yanow (1993/2005) connected organizational learning to organizational culture. Learning organization theory is relevant to many of the concepts related to program evaluation capacity (Cousins et al., 2008). Davidson (2005) noted, “a learning organization is one that acquires, creates, evaluates, and disseminates knowledge—and uses that knowledge to improve itself—more effectively than do most organizations. The best organizations tend to use . . . evaluations to build learning capacity” (Davidson, 2005, p. 3). Senge (1990) defined a learning organization as “an organization that is continually expanding its capacity to create its future . . . . But for a learning organization, ‘adaptive learning’ must be joined by ‘generative learning,’ that enhances our capacity to create” (p. 14). Furthermore, Senge posited five essential elements of learning organizations: (a) “personal mastery” focused on self-actualization through continued learning, (b) challenging ingrained “mental models” of long held beliefs, (c) “shared vision” for inspiration among people in an organization, (d) promotion of “team learning” through alignment of individual strengths for shared goals, and (e) “systems thinking” as the process to understand the whole by advancing the other four elements (Senge, 1990).

Program evaluation efforts provide valuable information for what Argyris (1999) called “double-loop learning” (p. 68). Essentially, double-loop learning is taking information from failures and changing variables to achieve success. In nonprofits, double-loop learning can occur from using program evaluation information to improve program processes, outcomes, or impact. Organizational learning occurs when the organization absorbs program evaluation data and makes positive changes based on the information (Preskill & Torres, 1999).The elements of learning organization theory that apply to the present study concern the degree to which nonprofit organizations seek evaluative information to use for reflective practice, to make necessary programmatic adjustments, and to further the organizational mission.

As previously mentioned in the program evaluation section, one of the main purposes for evaluation is formative. B. Johnson and Christensen (2008) simply defined formative evaluation as "evaluation focused on the evaluation object" (p. 588). Many of the evaluation approaches based on formative evaluation regard using information from the evaluation as a tool to facilitate improvement and learning within an organization (Patton, 1999). In fact, the Center for Disease Control considered organizational learning as a key component in developing evaluation capacity through promoting an environment focused on "lessons learned" rather than retribution for failures (Milstein et al., 2002). Patton (2010) made a similar point for the connection between the use of evaluations and learning organizations, "Social service organizations can improve effectiveness by becoming learning organizations" (p. 18).

Collaborations

Collaborative efforts among nonprofit organizations, funders, other stakeholders, and evaluators were described in the literature as a highly regarded strategy for building evaluation capacity in nonprofit organizations (Atkinson, Wilson, & Avula, 2005; Bozzo, 2002). In particular, developing collaborations among nonprofits, evaluators, and funders around common purposes and goals for evaluation has been identified as a feature needed to build evaluation capacity (Behrens & Kelly, 2008; Hendricks et al., 2008; Kegeles et al., 2005). However, authors noted that different disciplines involved in collaborative efforts often do not "speak the same language" or have the same values. Consequently, collaboration efforts should address these issues to help with communication (Milstein et al., 2002; Tang et al., 2002).

Duignan (2003) recommended collaborations among nonprofit organizations in order to present a united front regarding important evaluation agenda items for their specific sector. Related to Duignan's suggestion is the idea of agencies "pooling" their resources in order to build evaluation capacity (Bozzo, 2002). Stevenson et al. (2002) suggested collaboration for program evaluation efforts between agencies working on similar projects. In a parallel fashion, T. I. Miller et al. (2006) suggested "insourcing" (p. 83) as a way for nonprofit organizations to maintain a quality evaluation program by several nonprofits amalgamating resources to hire an expert evaluator. T. I. Miller et al. (2006) presented the concept as an alternative to evaluation capacity building strategies. Benefits noted include reduced costs to each individual organization, reduced amount of time devoted by program staff to the evaluation process, and the use of data from multiple organizations to evaluate larger community concerns.

Several authors noted the importance of nonprofits engaging in on-going collaborations with evaluation experts for technical assistance as a strategy to implement and maintain a quality evaluation agenda (Compton et al., 2002; King, 2002; Poole et al., 2000; Stevenson et al., 2002; Tang et al., 2002). There are different ways to create collaborations between evaluation experts and nonprofit organizations such as utilizing trained graduate students as technical experts (Arnold, 2006), hiring professional evaluators, or through United Way providing funds or direct support to their grantees (Hendricks et al., 2008),

Another crucial collaboration strategy for evaluation capacity building in the nonprofit sector centers on building partnerships with funders, particularly in relation to financial resources. Unfortunately, Carman (2007) found that although funders required evaluative efforts from nonprofit organizations, they rarely funded these evaluation activities. Yet funders were often the catalyst for program evaluation efforts in nonprofit organizations (Carman, 2010; Murray, 2005; Naccarella et al., 2007, Newcomer, 2001; Sowa et al., 2004; Urban & Trochim, 2009) but did not have a clear understanding of the lack of capacity within nonprofit organizations to institute quality evaluations (Kegeles et al., 2005).

Competing requests from multiple funders and other stakeholders tend to create obstacles and further burden the limited evaluation capacity in most nonprofit organizations (Carman, 2007; Ebrahim, 2010; Newcomer, 2001). Furthermore, research has suggested that funders are one of the least involved stakeholders with evaluation efforts in nonprofits (Cousins et al., 2008). It appears that a great divide between funding bodies' expectations and the ability of nonprofit organizations to meet these demands exists (Kegeles et al., 2005).

Reasons for this disconnect may center on the type of relationships between funders and providers. In particular, resource dependency theory and agency theory may help explain the relationships. Resource dependency theory recognizes that one group (providers) is beholden to another group (funders) for their existence. Providers are dependent on funders for resources so they use program evaluation strategies to enhance their efforts to acquire funding and other resources. Agency theory describes the roles between two parties as either principals or agents. In short, agents do the bidding of the principals, and trust between the two is sometimes lacking. The role of the principals is to ensure that agents are performing as expected (Carman, 2010, 2011; J. L. Miller, 2002).

Carman (2011) found that most nonprofit leaders have typically viewed their association with funding sources regarding evaluation efforts as scripted or "dictated." Related to "upward accountability" (Ebrahim, 2010) and agency theory, the premise is that funders dictate the evaluation agenda with little to no input from the nonprofit organizations. Consequently, nonprofit organizations often conduct evaluations to appease funders and often produce results just to retain funds. Carman argued that conducting evaluations "as simply an external accountability tool does an inherent disservice to nonprofit organizations" (p. 367). Ebrahim (2010) noted, "onerous data requirements can lead nonprofits to develop monitoring and evaluation systems that, although satisfying donor needs for information, are of limited value for internal learning and decision making" (p. 113). In this paradigm, open communication and mutuality of respect that is necessary for true partners is rarely present. Furthermore, these types of attitudes towards program evaluation do not facilitate collaborative efforts between funding sources and nonprofit organizations.

To reiterate the importance of funder involvement with program evaluation capacity in nonprofits, Carman (2007) concluded:

> Funders need to stop asking community-based organizations to provide them with reports designated for accountability purposes that simply monitor or report evaluation and performance data, and they need to start asking (and then rewarding) community-based organizations for reports designed to demonstrate how they are using evaluation and performance data to improve service delivery. (p. 72)

It is imperative that partnerships between nonprofit organizations and funding sources focus on multiple areas such as resource development, program improvement, and outcome attainment. A multifaceted approach is needed to produce a meaningful and quality evaluation agenda. Ultimately, both nonprofit providers and funders are stakeholders with an investment in advancing the mission of the nonprofit and making a positive impact in society.

Conceptual Framework

The conceptual framework for the present study includes five broad interrelated theories pivotal to understanding the role of and capacity for program evaluation in the nonprofit sector: mission centrality, accountability, program evaluation, evaluation capacity, and the significance of the collaboration between providers and funders in the nonprofit sector.

Mission centrality reflects the purpose for the existence of nonprofit organizations and should be the “North Star” for accountability efforts. First and foremost, legal mandates require a distinct mission in order to establish a nonprofit organization. The mission becomes the rallying point for the nonprofit leadership and the foundation for strategic planning (Axelrod, 2005; Brinckerhoff, 2000; W. A. Brown, 2005; Bryson, 2010; Herman, 2010; Knauft et al., 1991; Worth, 2009).

The legal nature of nonprofits, the community-investment in the organization, and the way that nonprofits secure resources (e.g., donations, grants, and government contracts) require oversight and accountability to multiple stakeholders. The review of the literature identified different theories of accountability. The commonality among these theories is the emphasis on a multi-focal approach which includes financial resources, human resources, service delivery, management practices, legal mandates, compliance to industry standards, and ethical practices of accountability (Candler & Dumont, 2010; Carman, 2010; Cutt & Murray, 2000; Ebrahim, 2010; R. S. Kaplan, 2001; Murray, 2005, 2010).

Program evaluation is one vital component of a comprehensive accountability strategy. Yet a review of the literature revealed that leaders of nonprofit organizations are often ignorant of program evaluation strategies (Carman, 2007). Consequently, a program evaluation theory that includes meaning, purposes, benefits, approaches, and utilization is an important element of the conceptual framework for the present study. The theoretical basis of program evaluation as a distinct discipline with standards of performance and the various approaches available for use inform the present study. Stufflebeam (2001, 2007), Stufflebeam and Wingate (2005), and Patton (1997, 1999, 2002a, 2002b, 2007, 2010, 2011) were primary sources for the description of program evaluation used in the present study. Information on the various categories and approaches of evaluation from these authors shaped my organization of the material. Patton’s emphasis on the utilization and flexibility of program evaluation in order to adapt to change is particularly salient to the needs of the nonprofit sector.

The works by Carman (2007, 2010, 2011) and Carman and Fredericks (2008, 2009) have been a predominant influence for evaluation capacity theory. Research by these authors found that nonprofit organizations often lack the capacity to conduct useful program evaluations. However, Carman and Frederick (2008) found that the majority of the nonprofits studied attempted to institute some level of evaluation despite the lack of capacity.

Finally, leaders of nonprofit organizations and their funder compatriots have impact on the role of and capacity for program evaluation in the nonprofit sector. As described in the review of the literature, collaboration between nonprofit organizations and funders has been a key strategy for building evaluation capacity in nonprofit organizations and advancing a program evaluation agenda. However, lack of trust, limited communication, unrealistic expectations, and perceived inequality between the two groups seemed to inhibit meaningful partnerships. The collaborative efforts (or lack thereof) between nonprofit service providers and funders have significant bearing on the quality of program evaluation in nonprofit organizations. An understanding between providers and funders regarding the role of and evaluation capacity for program evaluation in the nonprofit sector can be a catalyst for stronger collaborations.

Of these five points of the conceptual framework, program evaluation, program evaluation capacity, and collaboration between providers and funders are the focus of the present study (see Figure 1 for a visual representation of conceptual framework).

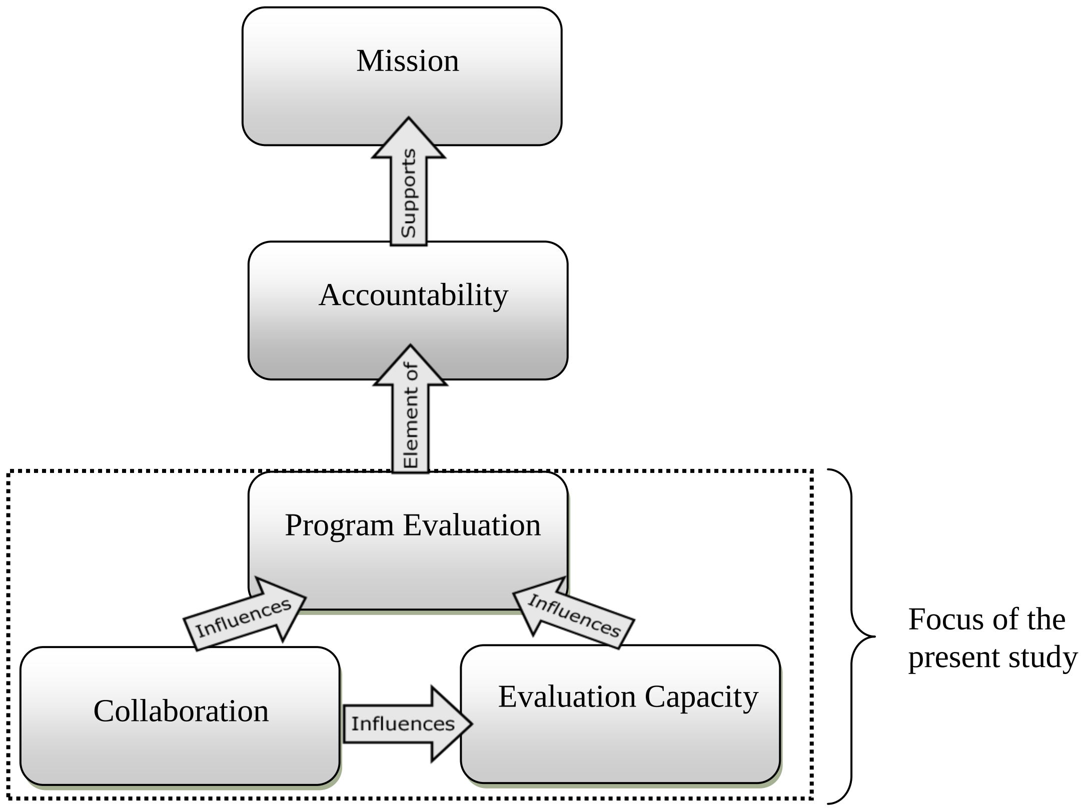

Figure 1: Conceptual Framework: A visual representation of the interrelationship of the different theoretical frameworks related to the role of and capacity for program evaluation in the nonprofit sector.

Summary

In summary, multiple stakeholders hold nonprofit organizations accountable for overall organizational efficacy, compliance with established standards, fiscal acuity, and evidence of alignment with the organizational mission. A key factor in the accountability paradigm is that these stakeholders require nonprofit organizations to provide empirical evidence of outcomes, program effectiveness, and overall impact of service delivery. Yet both nonprofit organizations and their funding partners have not invested sufficient resources into program evaluation capacity. Consequently, nonprofits continue to produce basic input and output information and erroneously tag these efforts as evaluation (Carman, 2007). In order for nonprofit managers to situate their organizations for the future, they must seriously invest in evaluation capacity

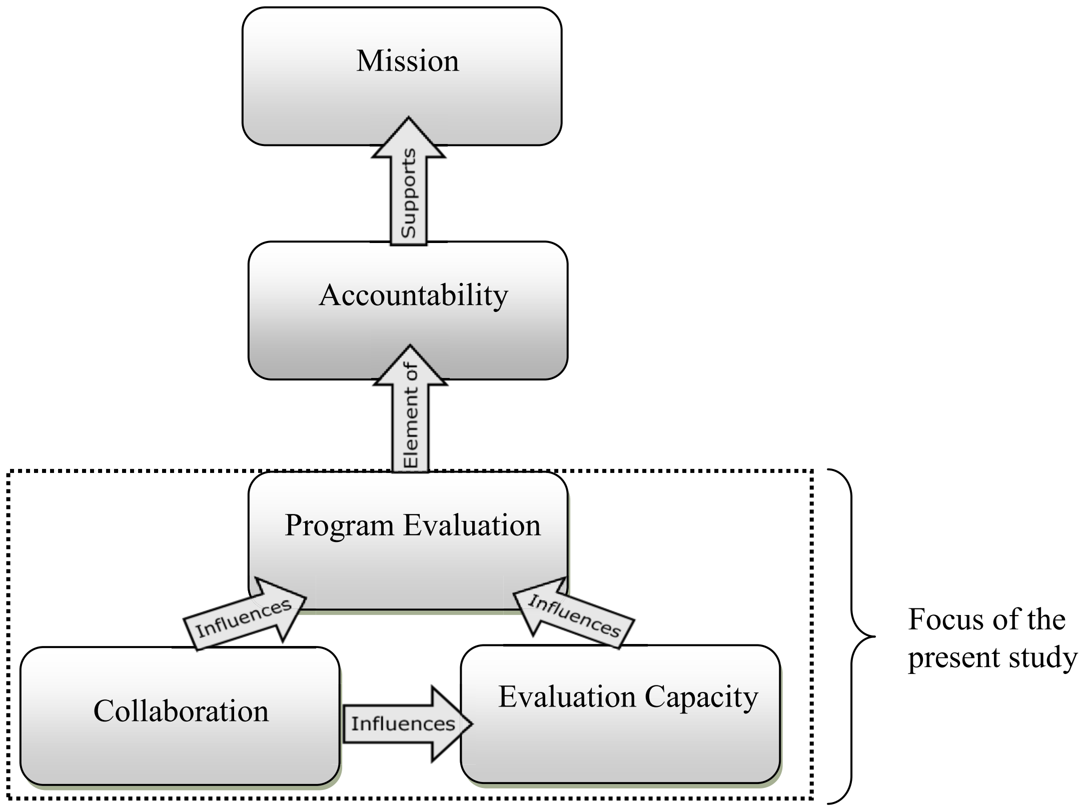

Figure 1: Conceptual Framework: A visual representation of the interrelationship of the different theoretical frameworks related to the role of and capacity for program evaluation in the nonprofit sector.

Summary

In summary, multiple stakeholders hold nonprofit organizations accountable for overall organizational efficacy, compliance with established standards, fiscal acuity, and evidence of alignment with the organizational mission. A key factor in the accountability paradigm is that these stakeholders require nonprofit organizations to provide empirical evidence of outcomes, program effectiveness, and overall impact of service delivery. Yet both nonprofit organizations and their funding partners have not invested sufficient resources into program evaluation capacity. Consequently, nonprofits continue to produce basic input and output information and erroneously tag these efforts as evaluation (Carman, 2007). In order for nonprofit managers to situate their organizations for the future, they must seriously invest in evaluation capacity building (Carman, 2007). Furthermore, funders must begin to "walk the talk" by providing specific funding for program evaluation efforts (Carman, 2007; Kehrer, 1993).

Clearly, the days of nonprofits "doing good" without providing evidence of their success are gone (Chaytor, MacDonald, & Melvin, 2002, p. 95). Nonprofit organizations are businesses that receive funding and other types of support from numerous sources. The demand from multiple stakeholders for accountability of program effectiveness, impact, and outcome attainment is strong. Additionally, information produced from evaluations provides opportunities for service delivery improvements, reflection, and creativity, which are necessary for a learning organization. However, the lack of evaluation capacity in nonprofit organizations creates a gap between what funders require and what nonprofit agencies deliver.

It is unclear as to what the perceptions of leaders of nonprofit providers and funding organizations are regarding the role of and capacity for program evaluation in nonprofit organizations. The present study investigated understanding these perceptions and the implications for development of program evaluation capacity. The following chapter presents the research design, the description of the setting, description of the participants, data collection methods, data analysis methods, ethical considerations, and the role of the researcher as instrument.

Chapter 3: Methodology

The present study was a two-part qualitative research design. The two phases of the present study were the Delphi surveys and in-depth interviews. The overall purpose of the present study was to understand the role of and capacity for program evaluation in the Northeast Florida nonprofit sector from the individual and collective perspectives of leaders of nonprofit human services providers and funding organizations. A review of the literature necessitated using a methodology that would provide a platform for leaders of nonprofit human services providers and funders to communicate their perceptions regarding these issues without fear of reprisal or loss of funding. The Delphi surveys method allowed for this anonymous communication between the groups. Additional information came from in-depth interviews.

Research Design

The purpose of the present study was to understand the perceptions of leaders from nonprofit human services providers and funding organizations regarding the role of and capacity for program evaluation in the Northeast Florida nonprofit sector. Gaining the subjective tacit knowledge of experts was a primary goal. Subjective knowledge from key informants is of great value. Patton (2002a) noted, "the perspective of others is meaningful, knowable, and able to be made explicit" (p. 341).

Marshall and Rossman (2006) noted that researchers commonly use qualitative methodologies when a purpose of a study is exploratory, descriptive, or explanatory. The present study solicited subjective knowledge and relied on the connoisseurship of the participating experts. Subjectivity and connoisseurship are key themes in qualitative inquiry

(Eisner, 1998; Patton, 2002a). Although Eisner (1998) promoted the idea of connoisseurship in educational criticism as a characteristic of the researcher, it was also relevant to the focus of the present study, as I used experts from nonprofit philanthropic and human services agencies as an essential element of the present study.

Other key factors of the present study involved providing a group of individuals a theater for the expression of voice and the promotion of understanding among collaborators. Eisner (1998) wrote about "voice" as an essential characteristic of qualitative research. The voices of participants are paramount to qualitative research. In essence, capturing the voice of participants creates deeper understanding of the topic. Moreover, it is important because it provides a vehicle to learn about experiences previously overlooked or discounted. The voice heard in qualitative research becomes a mode of increasing understanding and advancing empowerment (Marshall & Rossman, 2006). I provided opportunity for the expression of voice by leaders of nonprofit human services provider and funding organizations regarding the role of and capacity for program evaluation in the Northeast Florida nonprofit sector. This was important for providers as the relationship between the two groups often represents a power differential due to providers' dependence on funders for allocation of resources (Carman, 2011). Similarly, the present study provided an opportunity for leaders in funding organizations to give voice to their experiences regarding program evaluation and capacity issues.

Another aspect of the present study was the use of purposeful sampling, which is another characteristic of qualitative inquiry (Patton, 2002a). I approached specific people because of their leadership status in the Northeast Florida nonprofit sector. These leaders represented both providers and funders.

Delphi Surveys Method

The application of the Delphi method in the present study consisted of several survey rounds to a panel of experts. Experts in the present study were leaders of nonprofit human services providers (providers) and funding organizations (funders). Analytical data from each survey round became the building blocks for subsequent survey rounds. As a nominal group process, the Delphi method redacts the subjective knowledge of a panel of experts into collective quantitative information. Linstone and Turoff (2002) succinctly defined the technique “as a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem” (p. 3). The technique was “named after the ancient Greek oracle at Delphi, who offered visions of the future to those who sought advice" (Gupta & Clarke, 1996, p. 185). The Delphi method dates back to the first half of the 20th century at the Rand Corporation to forecast technological and military advances (Ekionea & Fillion, 2011; Geist, 2010; Gupta & Clarke, 1996; Hung et al., 2008; Landeta & Barrutia, 2011; Turoff & Hiltz, 1996).

In subsequent decades, researchers and practitioners used the Delphi method in other venues such as business, public administration, government, education, healthcare, and social services as a vehicle for curriculum development, consultation, decision making, business trends, planning, risk management, and for evaluative purposes (Adler & Ziglio, 1996; Armstrong, 2001; Clibbens, Walters, & Baird, 2012; Day, 2002; Hung et al., 2008; Landeta, 2006; Landeta & Barrutia, 2011; Linstone & Turoff, 2002). Lee et al. (2008) successfully employed the Delphi method for exploratory and descriptive research purposes focused on gaining perceptions of experts in the education evaluation profession. Karvonen, Ryynanen, and Kassi, (2009) used the Delphi method to gain information on the development of a certain type of computing. Ekionea and Fillion (2011) used the Delphi method to achieve their purpose of gathering information regarding good knowledge management practices. Although the use of the Delphi method continues, it was more widely used 20 to 30 years ago (Landeta, 2006).

Researchers developed different Delphi methods based on the purpose of a study (Hasson & Keeney, 2011; Hung et al., 2008). The classical or original Delphi, developed by Norman Dalkey and Olaf Helmer (1963) focused on building consensus of opinion from a group of experts. The authors initially explained the Delphi method:

> Its object is to obtain the most reliable consensus of opinion of a group of experts. It attempts to achieve this by a series of intensive questionnaires interspersed with controlled opinion feedback . . . . The technique employed involves the repeated individual questioning of the experts (by interview or questionnaire) and avoids direct confrontation of the experts with one another. (Dalkey & Helmer, 1963, p. 458)

Other common variations of the Delphi method include policy, decision, modified, argument, and disaggrative Delphi (Franklin & Hart, 2006; Hasson & Keeney, 2011; Hung et al, 2008; Tapio et al., 2011). The purpose of the policy Delphi is to generate understanding of a topic through different opinions of experts for which consensus is not a goal (Franklin & Hart, 2006; Turoff & Hiltz, 1996). The purpose of the decision Delphi is to produce a decision for a specified issue (Franklin & Hart, 2006; Hasson & Keeney, 2011). The focus of the modified Delphi is future forecasting or consensus building (Hasson & Keeney, 2011). The purpose of the argument Delphi is to ensure different opinions (Hasson & Keeney, 2011). The purpose of the disaggrative Delphi is to generate different thoughts in order to forecast different possible futures (Tapio et al., 2011). The commonality among these different techniques is the use of experts in a nominal group process (Hasson & Keeney, 2011; Hung et al., 2008; Linstone & Turoff, 2002;

Tapio et al., 2011)

I chose the policy Delphi variant for the present study. The policy Delphi variant is used when the aim of a study is to generate understanding rather than consensus from of a variety of expert opinions on the subject (Franklin & Hart, 2006; Turoff & Hiltz, 1996). The role of and capacity for program evaluation in nonprofits continues to be a topic with limited understanding. Through the present study, I provided new information regarding the role of and capacity for program evaluation from the perspective of providers and funders. Consensus was not a goal of the present study, but rather understanding the perspectives from leaders in both of these roles.

Linstone and Turoff (2002, p. 4) listed additional salient factors that indicated use of the Delphi method for the present study:

- Subjective information based on the connoisseurship of experts is sought.
- Representatives of the population under inquiry do not have a history of communicating. effectively on the subject and/or anonymity is needed to protect individuals or their organizations from political or other negative ramifications.
- Frequent group meetings are not possible.
- In order to promote diverse opinions through reducing the possible effects of individual or group pressure for conformity and other "groupthink"[1] characteristics.

The present study matched these indicators for use of the Delphi method. I wanted to learn from leaders in nonprofit human services provider and funding organizations regarding their perceptions of the role of and capacity for program evaluation in the Northeast Florida nonprofit sector. Furthermore, participation in the Delphi method was anonymous among participants. Consequently, providers gave input to funders without fear of reprisal. In addition, as a review

[1] Groupthink is the word originally described by Janus that "occurs when group members' desire to maintain good relations becomes more important than reaching a good decision. Instead of searching for a good answer, they search for an outcome that will preserve group harmony" (Levi, 2007, p. 156).

of the literature has demonstrated, communication between the two groups of leaders is limited so this method provides an opportunity to share information around a very salient topic. Additionally, the leaders who participated in the present study are very busy people. Their time is limited and meeting several times in a group would be difficult and would not allow for anonymity. Using the Internet for the Delphi surveys was a convenient way to conduct a nominal group process. Finally, without face-to-face real time interaction, participants were free to share their opinions without direct pressure for conformity from members of the panel.

In-Depth Interviews

The second phase of the present study included in-depth interviews with a subset of panelists and other leaders in the sector. Interview participants came from both providers and funders who also participated in the Delphi surveys. Interviews also included other leaders in the Northeast Florida nonprofit sector recommended because of their expertise or insight regarding the subject.

Interviews were used to gain deeper understanding regarding perceptions of the role of and capacity for program evaluation in the Northeast Florida nonprofit sector. First, interviews allowed two Delphi panelists with outlier responses to further explain or change their Delphi ratings. Another Delphi panelist (who was not an outlier) was interviewed to allow the opportunity to explore the topics in more depth. Second, interview questions (see Appendix B) explored the subject further with other leaders in the nonprofit sector as well as allowing for feedback from the findings of the Delphi survey. In addition, due to the qualitative focus of the present study, follow-up questions emerged during the course of the interviews.

Content Analysis

The hallmark of qualitative research is the pursuit of deep understanding through the subjective lens of the participants and the researcher. Humans engage in analysis of qualitative data regularly. Whether I observe the look in a loved one's eye, the smile on her face, or the tone of voice, I am interpreting data based on experience and my connoisseurship of the subject. Ultimately, data analysis is akin to the search for meaning. In a formal study, the researcher analyzes the qualitative data in a systematic fashion all the while acknowledging and controlling for subjectivity.

I primarily used content analysis with the data generated from the present study. Content analysis is one of the core components of qualitative research (Marshall & Rossman, 2006; Patton, 2002a) particularly for narrative qualitative data and interviews. Patton (2002a) defined content analysis as "qualitative data reduction and sense-making effort that takes a volume of qualitative material and attempts to identify core consistencies and meanings" (p. 453).

First, I organized the qualitative responses from the exploratory questions sent to panelists in the first Delphi survey round according to the questions, group membership, and individual. I used a combination of deductive and inductive (open, axial, and thematic) coding strategies for the data from the first Delphi survey round. B. Johnson and Christensen (2008) define inductive coding as "codes that are generated by a researcher by directly examining the data" (p. 539). Conversely, deductive coding uses pre-determined coding categories (Marshall & Rossman, 2006; Patton, 2002a). A review of the literature (Carman & Fredericks, 2008; Lee et al., 2008) provided additional information that gave me deductive coding strategies. These deductive coding strategies included ideas regarding the purposes of program evaluation such as program improvement, effectiveness determination, decision-making, strategic planning, programmatic design, outcomes assessment, and public relations. Through content analysis, I created categories and identified themes from the first Delphi survey round.

Analysis of the data from the interviews was similar to the steps used for analyzing the initial responses to the exploratory questionnaire. First, I organized the data by having the interviews transcribed from a digital recording. I read the transcriptions multiple times as well as listened to the recordings. Coding from the interviews was more deductive as I used the categories from the Delphi surveys as the foundation for the coding of the interviews. This is warranted because the interviewing data builds from the information from the Delphi surveys.

I also analyzed data from the Delphi surveys and interviews as a whole to determine similarities and differences. This also provided a broader perspective of the data. I wrote analytical and process memos as suggested by a review of the literature (B. Johnson & Christensen, 2008; Marshall & Rossman, 2006) as a way to document the analysis. B. Johnson and Christensen described the use of memos in qualitative data analysis as "reflective notes that researchers write to themselves about what they are leaning from their data" (p. 532). Each step involved a certain amount of interpretation based on the analysis (B. Johnson & Christensen, 2008). Ultimately, this analysis leads to data-informed interpretive conclusions on my part.

Research Questions

The three research questions posited for the present study were the following:

- What are the perceptions of nonprofit and funding organizations' leaders regarding the role (e.g., meaning, purposes, benefits, approaches, and utilization) of program evaluation in the Northeast Florida nonprofit sector?
- What are the perceptions of nonprofit and funding organizations' leaders regarding how the relationships between service providers and funders influence program evaluation efforts in the Northeast Florida nonprofit sector?
- What are the perceptions of nonprofit and funding organizations' leaders regarding the development of program evaluation capacity in the Northeast Florida nonprofit sector?

Description of Setting

The setting for the present study was the Northeast Florida geographical area. The areas commonly associated with this region include Baker, Clay, Duval, Flagler, Nassau, Putnam, St. Johns, and Flagler counties (Northeast Florida Regional Council, n.d.). However, for the present study, I included only Baker, Clay, Duval, Nassau, and St. Johns counties. One reason is that similar funding sources provide support for human services nonprofit organizations in these counties. Another reason is that the Northeast Florida Nonprofit Center designates these specific counties as their catchment area (Littlepage & Oldakowski, 2012). This was important as panelists needed to come from a specifically defined nonprofit sector.

A study commissioned by the Jessie Ball duPont Fund (Littlepage & Oldakowski, 2012) provided information on the nonprofit sector in this five-county region. At the time of the present study, the Northeast Florida nonprofit sector included over 1000 organizations with a 501(c) (3) IRS classification. Budgets for these nonprofits varied in size up to over $5 million. However, over 70% of nonprofits in Northeast Florida had budgets of less than $500,000. The predominant classification of nonprofit organizations was human services organizations. Most of the nonprofits in Northeast Florida were located in Duval County (70%).

I chose nonprofit human services providers and funding organizations from the Northeast Florida nonprofit sector for a few salient reasons. Primarily, participants needed to have the same contextual point of reference regarding accountability requirements, industry norms, evaluation efforts, and funding sources expectations. Next, my connoisseurship of the Northeast Florida nonprofit sector was an important factor in the present study that helped determine participants as well as provided a framework for discussion. Additionally, one of the research questions of the present study examined the perceptions of how the relationship between funders and providers influences program evaluation efforts. In order to address this question, participants must come from the same geographical area. As a result, participants knew that the information produced was directly relevant to their experience.

Participants

The present study was not dependent on a particular sample size because generalizability was not a goal. In fact, Donmoyer (1990) argued that the concept of generalizability was an outdated standard for many contemporary research designs such as found in education and the social sciences. Rather, the aim of the present study was to provide information that others might find useful and relevant to their practice or field of study (Eisner, 1998; Patton, 2002a). Lincoln and Guba (1985) labeled this use of qualitative data and findings as “transferability” (p. 124). Consequently, as indicated by a review of the literature, selecting the right participants for the present study was essential (Hung et al., 2008) and more important than the actual size of the sample. As Patton (2002a) eloquently stated, “while one cannot generalize from single cases or very small samples, one can learn from them-and learn a great deal, often opening up new territory for further research” (p. 46).

First, I delineated parameters for inclusion as panelists in the Delphi surveys phase of the study. The criteria involve inclusion in a least one of the following categories:

- Executive leadership (including board of directors) and upper management of nonprofit human services agencies and funding organizations preferably with a minimal of five years in the nonprofit sector.
- Personnel (paid or volunteer) of nonprofit human services organizations preferably with five or more years of experience in grant writing activities, program development, evaluation efforts, logic model construction, or program deliverables development.

- Personnel (paid or volunteer) of funding organizations, preferably with five or more years of experience with evaluation strategies, reviewing grants, allocation of resources, or grant compliance monitoring.

A review of the literature indicated other related requirements needed to participate in the present study such as having interest and time (Skulmoski et al., 2007). Participants also needed access to the Internet.

Second, I developed a list of potential panelists for the present study as suggested by a review of the literature (Clibbens et al., 2012; Ekionea & Fillion, 2011; Ferguson, 2000; Hasson & Keeney, 2011). In order to build the list, I used purposeful sampling and snowball sampling to choose the panel of eligible experts (Hung et al., 2008). Elements of snowballing sampling involved asking other experts in the nonprofit community for suggestions of who might be a good choice for representing the sector as it relates to this issue. Additionally, as a professional in the nonprofit world in the region under study, I have knowledge of and access to leaders of human services nonprofits and leaders in funding organizations. Based on this knowledge and the stated criteria, I generated a list of possible names and contact information to include on the panel.

Third, I contacted possible participants through calling on the phone, in-person, or through email to invite their participation in the present study. For the initial contact, I provided information on the parameters of the study. I did not reimburse panel members for participation. However, benefits included personal satisfaction of identification as an expert in their field and acknowledgement that their professional judgments matter (Landeta & Barrutia, 2011). Further benefits to participants include learning program evaluation issues through participation, increased understanding of different perspectives, and a summary of the results of the Delphi surveys (Gupta & Clarke, 1996; Landeta, 2006).

I obtained informed consent from those agreeing to participate in the present study. I secured signed informed consent for participation from those interviewed (see Appendix C). I obtained signatures on one copy and left an additional copy with them. I used electronic informed consents from those participating in the Delphi surveys phase of the study, as approved by IRB (see Appendix D). For an electronic informed consent, I had the informed consent narrative at the beginning of the survey. Participants had to check on the appropriate box indicating their agreement for participation before they were able to proceed with the Delphi surveys. This occurred with each round of the Delphi surveys. Participants were also able to print a copy of the electronic informed consent for their records.

In order to facilitate honest responses from participants, anonymity among participants is paramount. Consequently, participants on the panel knew the type of people involved (human services providers and funding leaders) and the fact that all participants were from the Northeast Florida region. However, participants did not know any other demographic or identifying information. Although anonymous to each other, I, as the researcher, knew the identities and responses of participants.

Data Collection

I used the Qualtrics software survey program to create the survey instruments in order to collect data for the Delphi surveys phase of the present study. I facilitated survey responses through the Internet. Additional data were collected through emails and interviews. I also used pre-determined questions with impromptu follow-up questions for the interviews phase of the study.

Delphi Surveys

In order to answer the research question that focused on the role of program evaluation, qualitative data were collected through an initial exploratory on-line survey that asked specific questions related to the meaning, purposes, benefits, approaches, and utilization of program evaluation. Similarly, the first survey also had questions related to the capacity for program evaluation. An additional question focused on how the relationship between providers and funders may influence program evaluation efforts (see Appendix D). Three other rounds followed the initial Delphi round where participants rated the different categories from the first round analysis. In all, four rounds of Delphi surveys were conducted for this phase of the present study.

First Delphi Survey Round. The initial exploratory survey included questions and areas of inquiry adapted from the literature review, most notably from Lee et al. (2008) as well as Carman and Fredericks (2009). Specifically, wording for questions one (Q1) and six (Q6) were adapted from Lee et al. (2008). Furthermore, the wording at the end of each question ("In a short phrase or sentence, please describe at least three [more are appreciated] thoughts that come to mind") was a direct adaptation from questions in the same study (see Appendix D). Additionally, questions two (Q2) and four (Q4) were also influenced by the information from Lee et al. (2008). Questions five (Q5) and seven (Q7) were adapted from the Carman and Fredericks (2009) study. Additionally, Q4 and Q6 were also influenced by the work of Carman and Fredericks (2009). Through email correspondence, the authors provided consent to use or adapt their questions for the present study (see Appendix E).

I contacted potential participants by phone or through email starting in late October of 2012. Criteria for participation as leaders in the Delphi surveys phase included staff or volunteers in human services nonprofits and philanthropic organizations with at least five years experience with a preference for individuals in executive leadership. In total, I initially attempted contact with 10 human services provider organizations' staff members (providers) and four philanthropic organizations' staff members (funders). Of the 10 providers initially contacted, eight agreed to participate in the study. The other two providers never returned my calls or emails. I subsequently learned that this particular human services provider organization, where both worked, was experiencing major leadership changes, which may account for the lack of response to my inquiry. Of the four funders initially contacted, three agreed to participate in the study and one returned my call and referred me to someone else in the organization. However, I did not follow-up on the referral as I had already secured a participant from that particular organization. Additionally, three more funders contacted me by email to participate in the study; another participant from a funding organization had referred them to me after I asked for referrals for the study.

On November 9, 2012, after an initial access glitch with the survey software program, I emailed the correct link for the first qualitative exploratory survey (see Appendix D) to the first group of leaders who agreed to participate in the study. As more joined the study, I also emailed them the survey link. After two weeks from the date I initially sent each person the survey link, I sent a reminder notice to those who had not completed the survey. I closed the survey on December 9, 2012, a month following activation of the survey. Of the eight providers, seven completed the initial survey (88%). Of the six funders, five completed the initial survey (83%). Only the 12 panelists who completed the initial exploratory survey had access to subsequent survey rounds. All 12 of these panelists remained with the process through the four different rounds.

Second Delphi Survey Round. The second Delphi survey round was a venue for panelists to rate the categories (i.e., role of program evaluation, essential program evaluation capacity elements) developed from the first round analysis. Initially, I was hesitant to have panelists rate the categories as I assumed that the panelists would rate the categories identically. Additionally, I felt it was important for the panelists to choose which categories were more important through a forced ranking system. I developed the survey as a forced ranked survey (see Appendix F) and emailed the link to panelists in the morning on January 5, 2013.

Within a few hours of sending the link of the survey, I realized that I had made a mistake in requiring panelists to force rank the categories. By using forced ranking, I was requiring panelists to rank categories when the panelists may perceive that some categories are equally important. Furthermore, I realized that by rating the categories, the means produced from the rated categories provided a way to rank order the information. This allowed panelists to respond to each category individually while providing information to determine group rankings. Additionally, a review of the literature indicated the use of rating with subsequent Delphi surveys rounds (Clibbens et al., 2012; Lee et al., 2008).

Consequently, I closed the survey the same day, sent emails to panelists explaining that there was an error in this survey, and recalled the link to the survey. I also informed panelists to expect a new link to the survey within a few days. However, one panelist did respond to the forced-ranked survey before I shut down access. This panelist was very gracious about the error and agreed to continue participation in the study.

Subsequently, I revised the survey for ratings responses. For the role of program evaluation in the Northeast Florida nonprofit sector, I asked panelist to rate on a five-point scale the level of importance of the different categories (1 = *not at all important*; 2 = *not very* *important;* 3 = *neither important nor unimportant;* 4 = *important;* 5 = *very important).* In the same manner, panelists were asked to rate their agreement regarding the essential program evaluation capacity elements relied on by the Northeast Florida nonprofit sector. A five-point scale was also *used* (1 = *strongly disagree,* 2 = *disagree;* 3 = *neither agree or disagree;* 4 = *agree;* 5 = *strongly agree*). The survey also contained a section for panelists to add comments. The thought was that this would allow feedback regarding their choices of ratings. I finally sent out the revised survey (see Appendix G) on January 9, 2013. I planned to allow for two weeks for everyone to complete the surveys. In the email I sent to panelists, I requested that they complete the surveys within a week. If a panelist did not complete the survey within a week, I sent a reminder. All panelists completed the survey within the two-week time period.

Third Delphi Survey Round. After the panelists completed the ratings for the second Delphi survey round, I computed the mean response for each category. Upon analyzing the results regarding the essential capacity elements required for program evaluation, I realized that I had asked the panelists to rate their agreement on whether program evaluation in the sector relied on the noted capacity categories. However, I did not ask them to rate to what degree they perceived the evaluation capacity categories existed in the Northeast Florida nonprofit sector. Consequently, for the third Delphi survey round, I took the program evaluation capacity elements categories that had a mean of four or higher (*agree* or *strongly agree*) and asked the panelists to rate to what degree these program capacity elements currently existed in the Northeast Florida nonprofit sector (see Appendix H). Panelists were provided a five-point scale to rate the degree the capacity elements are present (1 =*usually not present;* 2 = *often not present;* 3 = *sometimes present;* 4 = *often present;* 5 = *generally present*). A section for providing comments was also included at the end of the survey. I sent the link to the third survey on January 23, 2013. The same two-week time schedule with reminders sent at the half way point remained the target for completion. All panelists completed the surveys within the two-week period.

Fourth Delphi Survey Round. The fourth Delphi survey round was an opportunity for each panelist to review his or her individual ratings compared to the means of each subgroup affiliation (providers or funders) and for the combined group. Comments from previous surveys were included for their review as well. For this round, panelists had the opportunity to change individual ratings, justify individual ratings, or make any additional comments. In order to ensure the anonymity of panelists, each panelist received a separate survey with their scores and overall descriptive statistics for each category. Comments that had possible identifying information were not included or redacted to edit out the identifying information (see Appendix I for an example of a fourth Delphi round surveys). I asked panelists to complete the survey in a week. After a week, I sent a reminder to those who had not completed the survey. With the exception of one person, all panelists completed the survey within two weeks. The remaining panelist completed the survey within 16 days of release. I allowed extra time for the panelist to complete the survey because the two-week period ended on a weekend and because this particular panelist had several outlier responses. I felt the panelist's input on the last Delphi survey round was essential.

Interviews

The interview phase of the study commenced on February 28, 2013, two days after the last on-line Delphi survey was completed. The final interview was conducted nearly a month following the initial interview on March 21, 2013. I contacted potential interview participants by phone or via email to arrange interviews. If the person was unfamiliar with the study, I provided a brief introduction of myself and the purpose of the study. Every person agreed to an in-person interview. Prior to the interviews, participants received a copy of the "Interview Informed Consent" document. Participants signed the form prior to the interview and retained a copy for themselves. In particular, I chose two of the Delphi participants for interviews because their Delphi survey responses were outliers from the rest of the panelists. I interviewed another Delphi panelist recommended for an interview by a supervisor. I had contacted the supervisor, the chief executive officer of a funding organization, for an interview. However, the chief executive officer referred me to a senior executive who happened to be one of the panelists. It was important for me to include the perspective of this particular organization in more depth, so I interviewed this other panelist. This allowed the panelist to expound on the answers provided through the Delphi surveys. Other interview participants were leaders of philanthropic organizations or leaders in the Northeast Florida nonprofit sector recommended for interviews by other study participants. In one case, I first contacted the chief executive officer for an interview, but I received no response from my phone and email contact. I then contacted a senior executive who agreed to an interview.

Interview dates, times, and locations were arranged through email at the convenience of the interviewee. A structured set of questions (see Appendix B) provided a basis for the interviews; however, as each interview progressed, additional questions emerged based on the flow of the interview. For those leaders who did not participate in the Delphi surveys, I showed them the results of the survey towards the end of the interview after they answered all questions. I did this to get their feedback on the results.

I taped the interviews with two digital recorders, one as a backup in case the other failed. Additionally, I took notes that I later transferred to a word processing document. I employed a transcriber for the interviews and gave him the electronic files via a USB drive. The transcriber signed a confidentiality agreement (see Appendix J). The transcriber returned the transcripts via the USB drive. Electronic copies of the transcriptions were stored on the UNF SkyDrive. The interviews produced 310 minutes (just over five hours) of recorded conversation. The longest interview lasted 75 minutes and the shortest was 34 minutes. Additionally, transcripts from the recorded interviews generated approximately 155 double-spaced pages that resulted in an average of 26 pages per interview. The longest transcript was 36 pages while the shortest was 21 pages. Hard copies of the transcripts were stored in a safe at the home of the researcher.

Validity and Trustworthiness

As foundational concepts, I used the standards issued by the American Educational Research Association (AERA, 2006) for publication in any of their circulations to structure my research design for validity and trustworthiness. Two overall characteristics included the necessity of warrant and transparency. AERA defines warrant as "adequate evidence…provided to justify the results and conclusion" (p. 33). Transparency is "making explicit the logic of inquiry and activities" (p. 33) from the beginning to the end of the research project. Moreover, Howe and Eisenhart (1990) released their influential article on standards for qualitative research, which transformed the discussion of research rigor and methodological debates from a unitary positivistic-based debate to a broader adoption of standards applicable for both qualitative and quantitative research designs. The standards (pp. 6-8) included: (a) the fit with research questions, data collection, and analysis, (b) the effective application of specific data collection and analysis techniques, (c) alertness to and coherence of background assumptions, (d) overall warrant, and (e) value constraints. As previously noted, the research design is a fit for the research questions. Additionally, Kvale (1996) noted that this was one way of determining validity in qualitative studies. Furthermore, transparency regarding treatment of the data from collection to findings also increased the validity of qualitative studies. For example, I described in detail in Chapter 4 data analysis procedures to include independent critical reviewers of established categories. This also increased the validity of the present study.

Trustworthiness is a concept closely linked with the credibility of the researcher (Patton, 2002a). Patton described trustworthiness as "being balanced, fair, and conscientious in taking account of multiple perspectives, multiple interests, and multiple realities" (Patton, 2002a, p. 575). I chronicled efforts to understand the perspectives of the study participants through multiple surveys and interviews.

More important to increasing validity and trustworthiness of qualitative study is understanding the affect of the researcher as instrument on a study. I chose to emulate Patton's (2002a) stance of "empathetic neutrality" as a guide for my involvement in the present study. Empathetic neutrality "suggests that there is a middle ground between becoming too involved, which can cloud judgment, and remaining too distant, which can reduce understanding" (p. 50). To reinforce this concept, I am transparent about my own subjectivity as the researcher.

Researcher as Instrument

Transparency in a qualitative study is a key component, particularly as it relates to the researcher as an integral part of the study. Indeed, Patton (2002a) noted that the researcher is the main instrument of a qualitative study. Eisner (1998) also elucidated, "the self is the instrument that engages the situation and makes sense of it" (p. 34).

Married to the concept of the researcher as an instrument is the relevance of subjectivity in a qualitative research design (Eisner, 1998). In fact, the idea that any research is free of subjectivity is antiquated. Whether the design of the study is quantitative, qualitative, or mixed, the subjectivity of the researcher emerges from the genesis of the study to the final period on the last page of the written report. It is the researcher who chooses the topic based on subjective reasons, some known and some hidden within the sub-conscious mind. It is the researcher who chooses the questions to ask, the methodology to use, and the style of language of the written report (Eisner, 1998; Patton, 2002a). All of these are subjective and shape the research project. In qualitative research, researchers acknowledge and utilize subjectivity as a basic concept that requires transparency. Awareness of personal subjectivity is important. However, subjectivity is not negative but rather increases understanding through multiple perspectives and provides knowledge from an expert's point of view (Eisner, 1998).

The experience, knowledge, and skills of the researcher are elements of subjectivity. Eisner (1998) labeled this connoisseurship. Eisner defined connoisseurship as "the means through which we come to know the complexities, nuances, and subtleties of aspects of the world in which we have a special interest" (p. 68). Transparency regarding my professional connoisseurship as it related to the present study is important.

First, I am a social worker by profession and have worked in the nonprofit sector for over 25 years. Over 20 of those years were concentrated in supervisory positions. Currently, I am the Chief Operating Officer (COO) for a large nonprofit agency in Northeast Florida that provides services to victims of domestic violence and their families. I have been in this position for over 15 years, and one of my responsibilities includes creating, measuring, and reporting program outcomes to funders.

Over the years, I experienced frustration regarding the lack of organizational capacity (e.g., funds, knowledge, and time) to develop meaningful measures of program effectiveness. Moreover, I often felt that funders did not understand the pressures experienced by program staff to institute evaluative measures, and as a result, I sometimes resented the demands for outcome measures required by funders.

My initial interest in this research topic was to find a way to inform funders of nonprofit programs of these barriers and to find justification to remove or reduce funder requirements in this area. However, after a review of the literature, I realized my knowledge regarding evaluative practices was abysmal, and that my ignorance hindered my agency's ability to create an effective evaluation agenda. Eventually, I embraced the need for quality program evaluation in nonprofit organizations. I came to recognize the need for meaningful evaluative efforts that cultivate formative evaluation practices rather than the seemingly singular focus on summative evaluation. As a practitioner, I realized that instituting a quality evaluation agenda could improve service delivery, promote mission attainment, and add to the nonprofit sector knowledge base.

As I gained knowledge, I realized that other leaders in nonprofit organizations might experience these same limitations, but meaningful conversations or collaborations between nonprofit agencies and funding organizations regarding the subject are rare. This is not surprising. After all, funding organizations have the power to reallocate, reduce, or remove funds to nonprofit agencies. Accordingly, communication on this subject is often truncated. My hope was that I would advance communication on the subject between providers and funders in the nonprofit sector through this present study.

Obviously, I am not an objective observer of this subject. As a result, it was imperative that I was transparent and monitored my own subjectivity during the course of the present study. Peshkin (1988) noted researchers "should systematically identify their subjectivity through the course of their research" (p. 17). I increased rigor in this area by being constantly reflexive regarding how my perspective influenced the present study as suggested by Patton (2002a). Peshkin further suggested monitoring any intense feelings as indicators of subjectivity. Additionally, he proposed that the researcher uncover the different "I's" in life which come to light during the course of the study. For example, my known "I's" going into this present study were "I as COO" (of a human services nonprofit agency), "I as grant writer" (in my position as COO), "I as a board member" (of another local human services nonprofit), "I as grant reviewer" (of a local cooperative giving alliance), "I as donor" (to various nonprofit agencies and churches), and "I as consultant" (for local faith-based organizations). All of these personas had a vested stake in the present study, so it is important to identify and understand the possible influences on the present study.

I increased transparency in the present study and thus increased rigor by identifying myself to potential participants not only as a University of North Florida (UNF) doctoral student but also by my position in the community. I also was very cognizant of my interactions with participants during the interviews. I limited my input and comments during the interview phase of the present study. In doing so, I reduced the amount of exposure to my personal perspective on the subject, as my goal was to learn from the leaders I interviewed.

Ethical Considerations

Before collecting any data from participants, I submitted all required documents to the University of North Florida's Institutional Review Board (IRB). The IRB staff determined that the present study was exempt from further IRB review. The study was approved on October 26, 2012 (see Appendix K).

I took preemptive measures in the present study to protect the confidentiality of the individual participants and the organizations they represent. These steps were necessary to secure approval from the University of North Florida's (UNF) Institutional Review Board (IRB), which I needed to conduct the present study. As mentioned previously, I secured signed informed consent for participation from those interviewed (see Appendix C) and electronic informed consents (see Appendix D) from those participating in the Delphi surveys phase of the study as approved by IRB.

A primary consideration that I mentioned in a previous section is that I did not reveal the identity of participants in the Delphi surveys phase of the study to each other. Their interaction was solely through information of the particular group and combined group descriptive statistics and narrative feedback that was not identifiable. I also did not reveal the identity of the participants I interviewed to others.

Although participation was anonymous among the panelists, I knew their identity. Therefore, I did not use their real names, organizational affiliation, or any other identifying information. Instead, I used gender neutral pseudonyms or group affiliation for participants rather than his or her name. Any documents that have identifying information were maintained in my possession. I stored electronic documents on a UNF protected server (Osprey Skydrive). I maintained hard copies of documents in my home office in a safe and will shred them after the retention period prescribed by UNF's IRB.

After the interviews were completed, I downloaded the audio digital files and stored them on a secure server (UNF Osprey SkyDrive). Once transferred, I deleted the recordings from the recorder. I maintained digital recordings of the interview on the secure server until I completed the study, at which point, the digital recordings will be permanently deleted.

I used a transcriber in order to have written transcripts. I required him to sign a confidentiality agreement prohibiting him from disclosing any of the information to others. I provided the transcriber with an electronic copy of the interview. I delivered the recording in person through a portable hard drive (USB) to transfer to his hard drive for transcription purposes. The transcriber deleted any copies of the transcripts after returning the USB drive.

Delimitations and Limitations

The delimitations of the present study were as follows: (a) participants were selected from nonprofit human services providers and funding organizations from Northeast Florida, (b) I designated participants as leaders and/or "experts" in their respective fields based on pre-defined criteria, (c) there were a small number of participants, (d) participants were required to have access to and some familiarity with on-line surveys and Internet communication, and (e) the study occurred during a single point in time. A delimitation was also the use of purposeful sampling of the participants and criteria used for selection. These delimitations were necessary, as the design of the study was to obtain perceptions from a very specific group of individuals. Additionally, the participants had to come from the same geographical region because they needed to have a similar nonprofit sector context. Due to the qualitative design of the study, generalizability was not an intended outcome. As Patton (2002a) suggested, however limited the generalizability of a qualitative study is, information gleaned from a qualitative study may provide learning opportunities for others.

A core reason for choosing the Delphi surveys method for the nominal group process was that it provided anonymity among participants. Consequently, the delimitation of conducting the Delphi phase of the present study via the Internet was essential for the anonymity and confidentiality of participants. However, as noted previously, although anonymity existed among the participants, I, as the researcher, knew their identities and responses.

The present study was limited to the knowledge and experiences of the participants. A limitation was also the degree to which participants were candid in their responses. Furthermore, the extent of their specific knowledge or perceptions regarding quality program evaluation strategies influenced the input and conclusions of participants.

The culture paradigm of the Northeast Florida nonprofit sector influenced the present study. This presents a limitation to those particular customs, language, processes, laws, ethics, and perceptions held by the human services sub-sector of the nonprofit arena. Information regarding program evaluation comes from the experiences of human service provider agencies, which may differ from other nonprofit organizations that do not have similar pressures or external expectations as human service organizations.

Summary

In review, the research methodology used in the present study met the criteria established by Howe and Eisenhart (1990). First, the Delphi method was the most applicable fit for the present research inquiry. I sought knowledge from experts in the nonprofit human services field and funding organizations regarding the role of and capacity for program evaluation in the Northeast Florida nonprofit sector. Designed to cull and combine data from experts, the Delphi method provided a structured nominal group process to gather this information.

Second, the Delphi method provided a systematic process for gathering data needed for this inquiry. Additionally, I used a bricolage of qualitative analysis methods and data reduction techniques such as narrative analysis, coding, clustering, categorization, or classification strategies from the lens of my area of connoisseurship (Eisner, 1998; Marshall & Rossman, 2006; Patton, 2002a). Furthermore, I used basic descriptive analysis of ranking data through use of overall category means and standard deviations (Clibbens et al., 2012; Lee et al., 2008). The Delphi method is a structured process yet has a degree of flexibility for analysis of rounds such that both qualitative and quantitative analysis are effective analytical techniques.

The third criterion established by Howe and Eisenhart (1990) is "alertness to and coherence of background assumptions" (p. 7). This standard draws attention to the need for the researcher to be transparent about subjectivity. Both of these concerns have been addressed in length in a prior section of Chapter 3.

The fourth standard requires that a study exhibit overall warrant. Howe and Eisenhart (1990) described overall warrant as "responding to and balancing the first three standards" (p. 7). I demonstrated adherence to the first three standards as posited by Howe and Eisenhart. As discussed previously, the Delphi method was an appropriate fit for the research questions. Additionally, I used the data collected appropriately to justify any conclusions. Furthermore, it was important that I remained attuned to the impact of my own subjectivity and was willing to expand my preconceptions and theoretical base as data provided evidence of new information.

Lastly, the value of my inquiry is evident. The present study was transparent and conducted in an ethical manner. I explained each step of the process and explored any possible biases on my part. The study is also a value to the nonprofit sector. Based on the review of the literature, the ability of nonprofit organizations to implement and sustain quality program evaluation agendas is difficult. Therefore, information from the present study may help shape the future course of the role of and capacity for program evaluation in the nonprofit sector by providing information that may help advance necessary evaluation strategies in nonprofit organizations.

Chapter 4 provides the findings of the present study. The data analysis process steps are described in greater detail. Additionally, Chapter 4 has the presentation of the data based on each research question.

Chapter 4: Findings

The purpose of the present study was to understand the role of and capacity for program evaluation in the Northeast Florida nonprofit sector from the perspectives of leaders from nonprofit human services providers and funding/philanthropic organizations. The three research questions posited for the present study were the following:

- What are the perceptions of nonprofit and funding organizations' leaders regarding the role (e.g., meaning, purposes, benefits, approaches, and utilization) of program evaluation in the Northeast Florida nonprofit sector?
- What are the perceptions of nonprofit and funding organizations' leaders regarding how the relationships between service providers and funders influence program evaluation efforts in the Northeast Florida nonprofit sector?
- What are the perceptions of nonprofit and funding organizations' leaders regarding the development of program evaluation capacity in the Northeast Florida nonprofit sector?

The research design was exploratory in nature as the research questions centered on the perceptions of leaders from human services provider organizations (providers) and philanthropic/funding (funders) organizations in the Northeast Florida nonprofit sector. The structure of the research design focused on gathering this information through a nominal group process as well as through interviews.

Because of the exploratory nature of the present study, a qualitative study was the most appropriate design (Marshall & Rossman, 2006). To garner the initial qualitative data, I used the Delphi method for the nominal group process. Descriptive statistics from the Delphi surveys provided rating and ranking information to further understanding of the qualitative data produced from the initial exploratory Delphi round. The second phase of the present study included in-depth interviews that provided context and expanded insight of the topic. This combined research approach allowed for a multi-faceted examination of the subject that provided robust data to answer the research questions.

Description of Participants

The present study consisted of two phases conducted in the Northeast Florida nonprofit sector that includes Baker, Clay, Duval, Nassau, and St. Johns counties. The first phase was the Delphi multi-round on-line surveys. Panelists for the Delphi surveys were selected from provider and funding organizations. Delphi panelists were either recruited by me or referred from leaders in the Northeast Florida nonprofit sector.

The second phase of the present study centered on follow-up interviews with a select number of Delphi panelists and others. In particular, I selected two Delphi panelists that had outlier survey responses and another who did not have outlier responses. Additionally, leaders in other nonprofit organizations mentioned in the Delphi surveys or from interviews recommended from other study participants were also included in the interviews phase.

In total, between the Delphi surveys and interviews, 15 leaders representing 13 organizations participated in the study. Of these, nine (60%) were the executive leader (executive director, chief executive officer) for their organization. Five (33%) were senior executives in their organizations. The remaining participant (7%) was a staff member involved with program evaluation activities at a nonprofit agency.

Organizations

Information regarding each represented organization came from the participants, organizational websites, and GuideStar. Five were funding organizations, six were human services provider organizations, and two (representatives participating only in the interview phase of the study) were from other types of nonprofits. During the course of the study, one person worked at two different organizations. Consequently, I included demographic information from both of the organizations. Two of the organizations had two participating panel members each. In those cases where their answers regarding demographic information of the organization conflicted, I used auxiliary data such as agency websites and GuideStar to confirm the correct information.

Organizational age spanned from less than one year in operation to over 50 years. An equal number of organizations were in the 15-24 years and 25-50 years organizational age ranges. Two organizations have only been in existence for less than five years. (see Table 2).

Table 2

Organizational Age Frequencies

Age of Organizations	
50+ Years	2
25 - 50 Years	4
15 - 24 Years	4
5 – 14 Years	1
< 5	2
Total	13

Of the 13 represented organizations, only three had annual budgets greater than $10,000,000. The majority of represented organizations had an annual budget in the $1,000,000+ to $5,000,000 range. The organization with the smallest budget ($150,000-$349,999) was also the newest established organization represented in the study (see Table 3).

Table 3

Organizational Annual Budget Frequencies

Annual Budget of Organization	
$10,000,000+	3
$5,000,000+ – $10,000,000	0
$1,000,000+ – $5,000,000	4
$750,000 – $1,000,000	1
$500,000 – $749,000	1
$350,000 – $499,999	2
$150,000 – $349,999	1
Unavailable	1
Total	13

Overall, provider organizations had the largest staff size (50-99). Conversely, philanthropic organizations had predominately smaller staff sizes (see Table 4).

Table 4

Organizational Staff Size Frequencies

Staff Size of Organization	
50 - 99	4
20 - 49	2
10 - 19	3
1 - 9	4
Total	13

Organizations represented in the study provide a range of services to Northeast Florida, including at-risk youth services, services to victims of abuse, and support to families in crisis. Additionally, all five county areas (Baker, Clay, Duval, Nassau, and St. Johns) received services or financial support from at least one of the organizations represented in the study. However, only one organization was located outside of Duval County in St. Johns County.

Individual Participants

Fifteen leaders from the Northeast Florida nonprofit sector representing the 13 organizations noted previously participated in the Delphi surveys as panelists and/or interviewees in the study. Of the 12 Delphi survey panelists, three were also interviewed. Three other leaders (one from a philanthropic organization and two from other nonprofit agencies) only participated in the interviews.

The years of experience for study participants were extensive as 73% had more than 10 years experience working in the nonprofit sector. Only one provider panelist had less than five years experience in the nonprofit sector. Although this panelist did not meet the desired five years experience in the nonprofit sector, this panelist had over 10 years experience in human behavioral services at for-profit venues (See Table 5).

Table 5

Nonprofit Years Experience Frequencies of Study Participants

Years Experience	
10+ Years	11
6-9 Years	3
5 Years	0
2-4 Years	1
Total	15

All panelists in the Delphi surveys phase had at least a four-year college degree. However, a majority (60%) had advanced degrees (Juris Doctor, doctoral degrees, and master's degrees). The majority of the participants with advanced degrees were from the providers sub-group (See Table 6).

Table 6

Education Frequencies of Study Participants

Education	
Advanced Degrees (JD, MD, Doctoral, Masters)	9
4-year College Degree	6
Total	15

Age, race, and gender demographics of the total study participants (Delphi surveys and interviews) indicated more similarities than diversity except with the possibility of various age ranges. However, the majority of study participants (60%) were 45 years of age or older. Thirteen of the study participants reported being Caucasian/White. Gender was also not equally distributed with females representing 87% of the total group. Each represented sub-group (providers and funders) had one male participant (see Table 7).

Table 7

Demographic Frequencies of Study Participants

Categories		
Age	25 to 34	3
	35 to 44	2
	45 to 54	4
	55 to 64	4
	65 or over	1
	Unavailable	1
	Total	15
Race/Ethnicity	White/Caucasian	13
	Other	1
	Unavailable	1
	Total	15
Gender	Female	13
	Male	2
	Total	15

Summary of Methodology

The present study consisted of multi-round Delphi surveys and in-depth interviews. The Delphi surveys were the foundation of the study in order to gain initial information regarding the role of and capacity for program evaluation in the Northeast Florida nonprofit sector from the group of selected panelists. The first step of the Delphi method was selecting a panel of leaders through the criteria noted in Chapter 3. The second stage was the administration of the initial probing questionnaire (Ekionea & Fillion, 2011). The first questionnaire was exploratory and generative as panel participants responded to a few very broad questions with a list of ideas. Next, I analyzed and classified the qualitative information provided from the first Delphi survey round to create a list of ideas for rating in successive rounds. In ensuing rounds, panelists were asked to rate the list of categories generated. During analysis of each round, panelists had the opportunity to provide additional comments. However, all responses filtered through the researcher to protect anonymity among panelists as recommended (Dalkey & Helmer, 1963; Hung et al., 2008). The Delphi surveys consisted of four rounds.

The second phase of the present study focused on interviews with providers, funders, and other nonprofit leaders recommended because of their relevance to the subject. The interview phase of the study was critical to understanding outlier responses from the Delphi surveys phase as well as garnering information from other sector leaders. I analyzed and coded data from the interviews based on categories from the first Delphi survey results and reoccurring themes from the interviews. The interviews provided deeper and richer understanding of the role of and capacity for program evaluation in the Northeast Florida nonprofit sector.

Data Analysis

I used Marshall and Rossman’s (2006, p.156) procedures for analyzing the data as a guide for my own analysis. The analytic steps I employed were organization of the data, coding, data immersion, and constructing categories and themes. I employed these data analysis methods for both the Delphi surveys and the interviews. The Delphi surveys also produced quantitative data that I analyzed through using the overall means and standard deviations to rank the categories.

Delphi Surveys Analysis

The nucleus of data analysis for the Delphi method was content analysis of the narrative produced from this first round of open-ended questions. During the month the initial exploratory Delphi survey was active, I reviewed each panelist's responses after completion of the survey. This first step of data immersion provided the opportunity to assess patterns within individual responses. I also used this step to determine if panelists had asked any questions about the process for which they needed assistance. Additionally, prior to closing the survey, I began organizing and coding the data by writing each response for each question on a color-coded card. I coded the cards per question and by sub-group status (funder or provider). I also wrote the initials of each respondent on the back of the card. I did this in case I had a question about what I had written so that I could go back to the survey to verify the information if necessary. In some instances, a response may have generated more than one card as the response had multiple answers listed. For example, if the one response had independent thoughts typically joined by a conjunction, I wrote the two independent thoughts on separate cards.

Question 2 on the survey (Q2: What are the purposes for conducting program evaluations in nonprofit organizations?) was the first question I organized on color-coded cards. I chose this question to start the analysis as it directly related to the research purpose and the first research question. Q2 generated 29 responses from providers and 28 responses from funders. Using inductive analysis of the content of responses, I clustered words and ideas together that had similar themes. This analysis produced six initial categories (see Table 8). Each category was written on a white poster board and responses taped onto the poster boards.

The third question on the survey (Q3: What are the benefits to nonprofit organizations for conducting program evaluations?) was similar to the previous question (Q2) and was intended to elicit additional responses about the role of program evaluation. I documented and coded this group of responses as the next step in the analysis. Q3 resulted in 32 responses from providers and 19 responses from funders. The initial thematic clusters of words/ideas produced eight categories (see Table 8). Each category was written on a white poster board and responses taped onto the poster boards.

Table 8

Initial Content Analysis Categories for the Role of Program Evaluation

Q2	Q3	Combined Q2 and Q3
• Accomplishment of goals/reasons for program • Information, improvement, learning • Impact • Cost effectiveness/resource allocation decisions • Stakeholders • Alignment with mission	• Credibility and justification • Stakeholders information • Resource allocation/attainment • Promotes reflective practice • Identify negative stuff • Determine results • Accountability • Info for change	• Resource allocation • Stakeholders • Reflective practice • Info/improvement/ learning • Impact • Goal attainment/reason for program • Identify negative aspects • Change agent • Mission alignment • Justification/credibility • accountability

Note. Question 2 (Q2) was "What are the purposes for conducting program evaluation in nonprofit organizations?" Question 3 (Q3) was "What are the benefits to nonprofit organizations for conducting program evaluations?"

After I taped all of the responses for Q2 and Q3 on their respective poster board categories, I hung all of the poster boards in my study. This facilitated the ability to look at the mass of data (108 responses) at the same time (see Figure 2).

Color-coding the cards helped me to quickly identify the sub-group membership (providers or funders) of responders and note any patterns of responses based on group status.

This venue for analyzing the data also allowed me to immerse myself easily in the data just by standing in my study and thoroughly reading the responses multiple times in order to promote understanding. Marshall and Rossman (2006) labeled this immersion. I discovered emergent themes by reducing and combining the words/ideas clusters. I then redacted the categories for Q2 and Q3 to a single set of categories (see Table 8).

Figure 2: Color-Coding Content Analysis Process

The next step of data analysis included color-coding responses on cards from questions 1 (Q1: When you hear or read the phrase, program evaluation, what does this mean to you?) and 5 (Q5: How can nonprofit organizations use program evaluation information collected?). These two additional questions also generated responses related to the role of program evaluation in the Northeast Florida nonprofit sector. Q1 generated 24 responses from providers and 20 responses from funders. Q5 produced 30 responses from providers and 25 responses from funders. I assigned the color-coded responses to an already established category from the redacted/combined categories from Q2 and Q3. However, I parked responses that did not fit any of the categories until further analysis and subsequently used these responses to identify other categories or expand categories.

As I read and re-read the initial combined categories, I moved the different cards/responses to the most appropriate categories. With the mass of accumulated data related to the role of program evaluation (over 200 responses from Q1, Q2, Q3, Q5), I needed a way to organize the data that was reflective of the panelists' responses but contained within a manageable number of categories. A review of the literature, particularly Carman and Fredericks (2008) and Lee et al. (2008), provided ideas on how to combine and label categories related to program evaluation. These included ideas regarding the purposes of program evaluation such as program improvement, effectiveness determination, decision-making, strategic planning, programmatic design, outcomes assessment, and public relations. Using this information, I re-examined the clusters of words/ideas. I then developed new categories or combined categories on white poster boards after rearranging the color-coded cards into various clusters. Finally, I created categories by labeling each cluster of words/ideas as (a) *validate organizational credibility*, (b) *determine resources allocation*, (c) *inform program management decisions*, (d) *facilitate quality assurance*, (e) *assess impact*, (f) *advance organizational learning*, (g) *enhance communication with stakeholders*, and (h) *cultivate funding resources* (see Appendix L for words/ideas clusters with category labels regarding the role of program evaluation).

In order to provide context to Delphi panelists regarding the depth of meaning for each category, I used many of the words and ideas from their responses to define each category . The final eight categories with definitions regarding the role of program evaluation in the Northeast Florida nonprofit sector included the following:

- **Validate Organizational Credibility** through mission/vision alignment, accountability and transparency, justification of value/importance of organization, and strategic planning efforts.

- **Determine Resource Allocation** such as cost effectiveness measures, efficiency determination, budget and cost considerations, and asset management.

- **Inform Program Management Decisions** such as programmatic needs assessment, program design/development, and implementation.

- **Facilitate Quality Assurance** through verification of goals and objectives being met, program improvement strategies to include elimination/adjustments/corrections of negative/ineffective programs or strategies.

- **Assess Impact** such as demonstrating program effectiveness, success, improvement in lives of program participants, measurements of change.

- **Advance Organizational Learning** through reflective practice, feedback opportunities, education and training, increased knowledge, and direction for change.

- **Enhance Communication with Stakeholders** (staff, public, program participants, other organizations, policy makers) through marketing strategies, sharing of program results, and reciprocal feedback opportunities with program participants.

- **Cultivate Funding Collaborations** such as providing data regarding the efficacy of program(s), providing information learned from evaluations, providing information regarding program improvements/development, satisfying funder requirements, and a means for garnering financial support.

I used a critical reviewer to look at the eight final categories of program evaluation categories in relation to the responses for each category to see if the designation of responses with categories made sense. The critical reviewer for the role of program evaluation is a lawyer by profession who has worked for over two decades in nonprofit organizations. Additionally, she had experience in grant writing and was familiar with funding requirements regarding outputs, outcomes, and impact of programs.

I provided the critical reviewer with an electronic word processing file that had the responses from the first Delphi exploratory round that pertained to the role of program evaluation (see Appendix L). I had the words/ideas clustered in the different categories. I tasked the critical reviewer with reviewing each category of clustered words/ideas to critique if clusters seemed appropriate for the category as defined. I also informed her that if she did not think the grouping of words fit the designated category that we could discuss and, if necessary, get another person to weigh in on the categories if we could not come to an agreement.

She reviewed the information and, at first, questioned whether the categories “Enhance Communication with Stakeholders” and “Cultivate Funding Collaborations” might really be one category as she felt funders were stakeholders. However, after I provided information regarding the focus of the study as being the perceptions of the panelists rather than technical distinctions of definitions, she agreed that the panelists clearly distinguished funder stakeholders from other stakeholders. Therefore, she concurred that the categories aligned with the responses provided by the panelists.

Information regarding the development of program evaluation capacity in the Northeast Florida nonprofit sector was reflected in questions 6 through 8 (Q6, Q7, Q8). I recorded responses for each question on color-coded cards in the same manner as previously stated. Question six (Q6: What are the main challenges the nonprofit sector in Northeast Florida faces regarding program evaluation?) provided 23 responses from providers and 20 responses from funders. Question 7 (Q7: What [if any] are the resources needed to develop capacity in order to sustain program evaluation efforts in the Northeast Florida nonprofit sector?) produced17 provider responses and 18 funder responses. Finally, question 8 (Q8: How do funders and providers in the Northeast Florida nonprofit sector work together to develop program evaluation strategies [e.g., objectives, outcomes, evaluation purposes, uses of evaluation results, evaluation approaches, capacity issues, resource allocations]?), prompted 17 responses from providers and 18 responses from funders.

Analysis for the capacity for program evaluation followed the same steps as mentioned previously (organization, color-coded cards, immersion in data, analysis of words/ideas clusters for themes, category development, reassessment of clusters, and use of white poster boards with initial and redacted categories). Q6 was a negative response question (challenges regarding capacity) and the responses clustered to show a needed capacity element. For example, if a panelist noted that not having enough money was a challenge, then I clustered the response with other responses regarding the need for financial resources as an element of capacity. Responses to Q8 focused on capacity building issues as well as the relationship between providers and funders. Consequently, although not intended from the design of the survey, responses to question 8 were clustered with responses related to capacity.

Responses related to the capacity for program evaluation were approximately half of the responses for questions related to the role of program evaluation. Answers were more succinct and similar. Initial words/ideas clusters that came from the analyzed data had themes of time, money, utilization of experts, designated and skilled nonprofit staff, culture, unreasonable expectations, evaluation design knowledge, collaboration, training/knowledge, lack of knowledge, and client participation. Through further content analysis, I combined, reduced and created categories by labeling the clusters. The labeled clusters for essential capacity elements were (a) *sufficient time*, (b) *sufficient financial resources*, (c) *sufficient human resources,* (d) *a positive culture*, (e) *realistic expectations*, (f) *functional program evaluation designs/methods*, (g) *ongoing collaboration*, and (h) *ongoing training* (see Appendix M for words/ideas clusters with category labels regarding essential capacity elements).

Similar to the roles of program evaluation categories, I used many of the words and ideas from the Delphi panelists' responses to define each category related to needed capacity elements.

The final eight categories with definitions regarding the needed capacity elements for program evaluation in the Northeast Florida nonprofit sector included the following:

- **Sufficient Time** available for staff to plan, implement, analyze, and reflect on results from program evaluation.
- **Sufficient Financial Resources** specifically designated for program evaluation efforts.
- **Sufficient Human Resources** such as skilled and designated evaluation staff or access to/partnerships with experts (e.g., professional evaluators, researcher community, program evaluation technical assistance consultants)
- **A Positive Culture** (organizational and sector-wide) that advances program evaluation efforts such as prioritizing resources, willingness to accept feedback, openness to change, and a readiness to reflect and act on results.
- **Realistic Expectations** from the philanthropic community regarding program evaluation efforts such as evaluation design, expected outcomes, reporting requirements, and funders responses to the results.
- **Functional Program Evaluation Designs/Methods** that take into account feasibility, appropriate measurement tools, access to data, infrastructure to accommodate data, and issues related to engaging program participants in the evaluative process.
- **Ongoing Collaboration** in and between the philanthropic and provider communities through negotiating appropriate evaluation strategies, sharing resources, providing results, and opportunities for discussion.
- **Ongoing Training** for providers and funders regarding program evaluation strategies/best practices and available resources.

As with the section on the role of program evaluation, I engaged another critical reviewer to determine if the categories and responses were in alignment. The critical reviewer for this section received a doctorate in educational leadership and works in a higher education setting. Similar to the tasks provided for first critical reviewer, I provided the second critical reviewer with an electronic word processing file that had the responses from the first Delphi exploratory round that pertained to capacity regarding program evaluation (see Appendix M). I had the words/ideas clustered in different categories. I provided an overview regarding the purpose of the study and the questions related to program evaluation capacity. I tasked the critical reviewer with reviewing each category of clustered words/ideas to critique for appropriate fit with the category as defined. I provided her my phone number in case she had any questions or needed clarification. After reviewing the data, she concurred that the responses and categories were in alignment and did not have any suggestions for improvement.

Question 4 (Q4: What are some program evaluation approaches used in the Northeast Florida nonprofit sector?) produced 24 responses from providers and 15 responses from funders. Approximately 11 other responses from other questions focused on the definition of program evaluations and the methods used. I did not use these responses in the analysis for the categories in the subsequent Delphi survey rounds. Rather the information provided understanding of the knowledge panelists have regarding program evaluation as a whole and findings for the first research question.

In total, the first exploratory Delphi survey round provided over 300 responses for analysis regarding the role of and capacity for program evaluation. An analysis of these data produced eight categories each related to the role of and capacity for program evaluation, which became the foundation for the subsequent Delphi survey rounds. As mentioned previously, the remaining 39 responses for Q4 and the 11 responses related to program evaluation methods and definitions provided additional data for answering the research questions. With 370 responses from the 12 panelists, the exploratory Delphi survey round generated rich data that took approximately a month to analyze.

Results from the second, third, and fourth Delphi survey rounds produced descriptive statistics and qualitative narrative. I organized the quantitative data on Excel spreadsheets and calculated the frequencies, means, and standard deviations in each category for each round. I used the overall means of each category of each section to rank the findings.

Interviews Analysis

The first step of analysis for interviews was taking the transcripts and organizing the data. In order to do this and make referencing easier, I numbered each line and each page of every transcript. Secondly, I followed Patton's (2002a) process of "developing some manageable classification or coding scheme" (p. 463) when analyzing transcripts. I established the following initial categories for coding transcription text:

- The categories from the "Role of Program Evaluation" from the Delphi surveys
- The categories from the "Program Evaluation Capacity Elements" from the Delphi surveys
- Relationship of providers and funders
- Development of program evaluation capacity
- Reflective Practice

I chose these categories based on the results from the Delphi surveys, the research questions from the present study, and a recurring topic (reflective practice) from the interviews.

For the next step in data analysis of the transcripts, I checked each transcription for accuracy by listening to each recording while reviewing the document and corrected a few minor errors in spelling and words. After I completed the review for accuracy, I immersed myself in the data by reading each of the transcripts again. During the subsequent readings of the transcripts, I used the categories mentioned previously to begin to code narrative. I reviewed each transcript at least five times.

Further analysis of the interview data included using content analysis to discover repetitive words, ideas and common themes (see Figure 3 for a word cloud of some repetitive words and ideas from transcribed interviews). Further inductive analysis of the identified ideas from the coded data unearthed themes and information to answer the research questions (Patton, 2002a).

Figure 3. The sizes of words in this word cloud are relative to the frequency of repetition in the transcriptions as compared to the other words/ideas from the transcriptions.

Presentation of the Data

I present the findings from the present study per research question. The primary findings are a combination of qualitative data gleaned and analyzed from the Delphi surveys and the in-depth interviews. Descriptive statistics from the final Delphi round provide ancillary information regarding rating and ranking of the different categories. Data from the interviews reinforced the findings from the Delphi surveys. In fact, each interview participant who did not participate in the Delphi surveys agreed that the results of the Delphi surveys represented the role of and capacity for program evaluation in the Northeast Florida nonprofit sector. Participants in the interviews also provided detailed observations regarding the different categories regarding the role of program evaluation in the Northeast Florida nonprofit sector.

The focus of the research questions was on the perceptions of the leaders from nonprofit human services providers (providers) and funding (funders) organizations. A hallmark of qualitative research is the importance of expressing the "voice" of study participants (Eisner, 1998). Therefore, I chose to use excerpts from interviews in the presentation of the data in order to convey participants' insights and points of view on the subject.

The leaders who participated in the Delphi surveys and the interviews have an elite status in the Northeast Florida nonprofit sector. In order to maintain their confidentiality, I used gender neutral pseudonyms for the study participants. Furthermore, the description of their positions and descriptions of the affiliated organizations are intentionally vague.

Research Question #1

The first research question put forward centered on the role of program evaluation from the perceptions of nonprofit human services and funding organizations' leaders in the Northeast Florida nonprofit sector. Aspects under consideration included the meaning, purposes, benefits, approaches, and utilization of program evaluation. Initial findings that answer this question came from the results of the exploratory Delphi survey phase of the study. Interviews provided supporting data and rich perspectives regarding the role of program evaluation that mirrored the results from the Delphi surveys.

Definitions and approaches. How leaders of nonprofit and funding organizations define the meaning of program evaluation is foundational to their perception of the role of program evaluation. Leaders often did not readily identify some efforts as evaluation. For example, one executive leader of a funding organization was reluctant to define an assessment of impact as a program evaluation strategy:

> It wasn't evaluation, what it was, was trying to help organizations that provide services figure out how to look at themselves, look at customers, clients, patients, whatever you call folks who benefit from the services of those organizations, and see if they really were being helpful. I guess is the simple way to put it. And if they weren't being as helpful as they thought they were being, how could they be, how could they reinvent themselves almost?

Yet the leader described a program evaluation approach implemented by organizations that helped nonprofit providers assess impact and areas for improvement of service delivery. The possible disconnect may be how leaders define program evaluation. This is not surprising as responses from the first exploratory Delphi survey regarding the meaning of program evaluation were divergent and focused more on (a) the different uses or benefits of program evaluation or (b) broad program evaluation designs. Some examples of panelists' responses were the following:

- "Combination of quantitative metrics and qualitative measures in the evaluation" (Funder Delphi panelist)
- "Formative and summative" (Funder Delphi panelist)
- "Tedious and difficult data collection" (Provider Delphi panelist)

However, three Delphi panelists provided broader definitions with focus on the process of program evaluation.

- "Method to collect, analyze and use information on the effectiveness and efficiency of a program or project." (Provider Delphi Panelist)
- "The organization has adopted a methodology for measuring and/or tracking program objectives." (Funder Delphi Panelist)
- "Understanding the process involved that shows a direct correlation between the outcomes expected and the activities that are supposed to help lead to these outcomes." (Funder Delphi Panelist)

Delphi panelists noted several program evaluation approaches used in the Northeast Florida sector. The responses represented a range of knowledge regarding program evaluation approaches. Many of the responses focused on the actual instruments (e.g., surveys, journals, and tests) used for the evaluation. Other responses focused on different models (e.g., empowerment, logic model, outcomes, accreditation, and reflective practice). Some responses focused on evaluation designs or methods (e.g., longitudinal, quasi-experimental, focus groups, and interviews). However, other responses were not consistent with program evaluation strategies (e.g., CEO search surveys and board self-evaluations). Two panelists (a funder and a provider) concluded that organizational-wide external evaluations were preferred over internal programmatic evaluations.

Delphi survey results. The first exploratory round of the Delphi surveys consisted of eight qualitative questions of which half of the questions (Q1, Q2, Q3, Q5) had responses related to the role of program evaluation. Eight closely related yet distinct categories regarding the role of program evaluation were identified through an analysis of the over 200 responses to these questions. Subsequent rounds produced descriptive data from providers and funders regarding the level of importance from 1 (*not at all important*) to 5 (*very important*) of each of these roles of program evaluation (see Table 9).

I ranked the findings according to the means of each category from all of the panelists (providers and funders) ratings. As noted in Table 10, seven of the eight categories regarding the role of program evaluation had mean ratings higher than 4. The category with the highest mean score was assessing the impact of programs. This particular category also had the least variation of ratings among panelists as indicated by the standard deviation of .39. The lowest rated category (validate organizational credibility) had an overall mean of 3.92 and the greatest variation with a .79 standard deviation. In fact, one panelist from each sub-group (providers and funders) rated this at a 3 or lower which lowered the overall mean. The difference between the

Table 9

Program Evaluation Rating Responses Frequencies

Categories	Rating Responses				
	1	**2**	**3**	**4**	**5**
Assess Impact of Program	0	0	0	2	10
Advance Organizational Learning	0	0	0	5	7
Cultivate Funding Collaborations	0	0	0	7	5
Inform Program Management Decisions	0	0	0	8	4
Enhance Communication with Multiple Stakeholders	0	0	0	8	4
Facilitate Quality Assurance	0	0	1	7	4
Determine Resource Allocation	0	0	2	5	5
Validate Organizational Credibility	0	1	1	8	2

Note. The anchors of the scale were 1= *not at all important*; 2 = *not very important*; 3 = *neither important nor unimportant*; 4 = *important*; 5 = *very important*

highest rank category and the least was less than one (.91).

The ranking of the overall means per subgroup of providers and funders had similarities and differences. The first ranked category (assess impact of program) was ranked number one for both providers and funders. Similarly, the category with the overall lowest mean (validate organizational credibility) was also ranked the lowest for each subgroup. However, the rankings of the six other categories were different per subgroup status. The second and third place rankings for providers focused on using program evaluation as a means to cultivate funding, and provide information to stakeholders. The second and third place rankings for funders focused on informing program management decisions and facilitating quality assurance. Additionally, both had advance organizational learning in the top three rankings (see Table 10 for detailed information).

Table 10

The Role of Program Evaluation in the Northeast Florida Nonprofit Sector Delphi Survey Results

Categories	**Combined *M* (*SD*)**	**Funders *M* (*SD*)**	**Providers *M* (*SD*)**
Assess Impact of Program	4.83 (.39)	5.00 (0.00)	4.71 (.49)
Advance Organizational Learning	4.58 (.51)	4.80 (.45)	4.43 (.53)
Cultivate Funding Collaborations	4.42 (.51)	4.00 (0.00)	4.71 (.49)
Inform Program Management Decisions	4.33 (.49)	4.40 (.55)	4.29 (.49)
Enhance Communication with Multiple Stakeholders	4.33 (.49)	4.20 (.45)	4.43 (.53)
Facilitate Quality Assurance	4.25 (.62)	4.40 (.55)	4.14 (.69)
Determine Resource Allocation	4.25 (.75)	4.40 (.55)	4.14 (.90)
Validate Organizational Credibility	3.92 (.79)	4.00 (1.22)	3.86 (.38)

Note. In those instances where the mean is the same but the standard deviations are different, I ranked the category with the lower standard deviation higher.

These results from the Delphi surveys clearly indicated distinct roles of program evaluation categories between the highest ranked category (assess impact of program) and the lowest ranked category (validate organizational credibility). Yet leaders interviewed did not make these types of distinctions. After analyzing the interview data, I identified four major areas of interest regarding program evaluation: (a) impact and mission validation, (b) decision-making tool, (c) reflective practice, and (d) communication with stakeholders.

Impact and mission validation. An analysis of the interviews suggested that two of the categories regarding the role of program evaluation (assess impact and validate organizational credibility) from the Delphi survey results were interrelated. The reason was that leaders interviewed linked the impact (success, outcomes/change, effectiveness) of a program to organizational mission.

All of the leaders interviewed discussed the need for providers to determine the impact (success, outcomes/change, effectiveness) of programs. Jessie Franklin, senior executive with a provider nonprofit, noted the need for nonprofits to address their impact in the community:

> What we are always trying to come back to is making a change in the community. Are you affecting a social issue that needs to be addressed in the community? So not just, how many or how much you're doing, but what is the impact of what you're doing?

Taylor Green, a senior executive with a provider nonprofit, also viewed assessing the impact of a program as a key role of program evaluation. Moreover, Taylor expressed concern for nonprofits that do not implement program evaluation strategies:

> I was really struck by how many people were saying [at a meeting by a funder requiring evaluative measures] that they just didn't have time to do this. And I thought, "Why aren't you doing it already? How can you not be evaluating your work in some way?" I find it hard to believe that there wasn't some way of determining if they were being successful or making a difference or if they were being efficient and effective and all of those things.

Jordan Duncan, who was only involved with the interview phase of the study, also found assessment of impact to be a key role of program evaluation. Jordan, a senior executive with a funding organization, has many years experience in the nonprofit sector. Jordan described the growing emphasis for nonprofit organizations to provide evidence of impact. Jordan remarked, "I think in today's funding environment it's critically important that programs look at the success and impact, you know, being able to measure if they are making a difference."

The emphasis from funding organizations on using program evaluation strategies to assess impact is not lost on human services providers. Taylor Green, a senior executive with a provider nonprofit, noted the emphasis placed on determining impact:

> I will credit [name of funder] and donors [for] understanding that you have to show that you really are making a difference and that you're not just nice but necessary, and that it mattered. . . . If you're not being able to speak about how things are making a difference, then you're a charity and who cares?

Indeed, this emphasis from funders on using program evaluation to determine impact can also have negative consequences for some provider nonprofits. Morgan Evans, a senior executive with a funding organization, seemed keenly aware of this reality. Morgan remarked, "We do decrease funding when we see that there's not impact and we have defunded some organizations for lack of impact… And that's going to continue."

Pat Baur, a senior executive with a provider nonprofit, was a strong promoter of the use of program evaluation to advance the mission of an organization. Although Pat did not participate in the Delphi phase of the study, Pat was receptive to an interview. Pat expressed great interest in the study on program evaluation. Pat was not surprised that determining the impact of an organization was the highest ranked category from the Delphi surveys:

> I really think that is so reflective of the culture of this community, because people see an evaluation after the fact, "we've done the programming, what is the impact?" As opposed to the tool to get there. I see that evaluation is my tool for getting there. But the evaluation is to make sure that I have this alignment with our vision, our mission, where we are headed, what's the roadmap. And the evaluation keeps me informed on that.

Other leaders interviewed indicated that program evaluation should be a continuous element of strategic planning. Jordan Duncan, a senior executive with a funding organization, remarked, "I think that one challenge is that a lot of people evaluate in the end, after something is over, when it really should be an ongoing process." The idea of using program evaluation as a strategic tool was reinforced by Jessie Franklin, a senior executive with a provider nonprofit, who remarked, "the benefit [of program evaluation] is that organizations have to think about what they're trying to accomplish before they start doing the work." Taylor Green, a senior executive with a provider nonprofit, also concluded that program evaluation is useful for strategic planning and mission alignment:

> I think that's got to start with your board too and throughout the whole strategic planning process. If you really have a process by which you say, "Here's our mission, and here's our strategy, and here's where we're going to spend our time and energy…." If we can't show that we're getting there than how do you know that you're going towards your mission.

Pat Baur, a senior executive with a provider nonprofit, also affirmed the idea that program evaluation should be embedded in organizational strategic planning and is the genesis for quality program development. Pat remarked, "I always start with 'What is the evaluation?' If you look at our strategic plan, you'll see the main metrics, and so we really use evaluation as very much a foundation."

Decision-making tool. Leaders in the Northeast Florida nonprofit sector interviewed also described the role of program evaluation as a way to inform program management decisions (rated at 4.33 with a fourth place ranking on Delphi surveys), to facilitate quality assurance (rated 4.25 and ranked fifth on Delphi surveys), and to determine resource allocation (rated 4.25 and ranked sixth on Delphi surveys) . Jordan Duncan, a senior executive with a funding organization, encapsulated these roles of program evaluation as a means to make "course corrections." During the interview, Jordan used words such as "tweak," "change tracks," "do something different," or "something that needs to be changed" to emphasize that formative program evaluation is essential to program delivery. Jordan described the role of program evaluation as an on-going process rather than an end product:

> You don't want to ask somebody to evaluate something just for the sake of evaluating it, but what is the benefit to the clients, to the customers that are being served, and how will it help inform your process? So how will you take what you learned from that and use it?" You can't just evaluate things at the end, you have to evaluate or document as you're

> moving along, because if you don't apply what you're learning to what you're doing as you're doing it, then really the impact of that grant, or the success of that grant, or the success of the people you're trying to help, is really affected. I think that one challenge is that a lot of people evaluate in the end, after something is over, when it really should be an ongoing process.

Jordan also concluded that formative program evaluation approaches were particularly important for pilot projects or when an organization is trying out different program designs. In both of these cases, managers need data in real time to improve, adjust, revamp, or eliminate programs. In the same manner, Taylor Green noted using data collected the first half of the school year to make changes during the second half of the school year.

Morgan Evans, a senior executive with a funding organization, agreed that using program evaluation information to improve services is essential. Morgan is a self-described "numbers" person. However, Morgan was quick to acknowledge that both quantitative and qualitative data are necessary for providing information for programmatic change. Morgan provided a poignant example of how program evaluation efforts produced information that led to immediate changes within the program:

> Through those focus groups, then we were able to identify, and really, honestly, it was the qualitative portion that gave us more food for thought for organizational improvement than the quantitative. So, one of the things that was brought up was, when a therapist comes to counsel a student, they're doing it at school, they get to meet with them at school, and so, often, there is an announcement made over the loudspeaker, "Johnny your counselor's here." And so, you know, it's the students in the room, with that student, hearing that Johnny's going to a counselor. Well that's, number one, it's a violation of their privacy. And, number two; it's a disincentive toward counseling. That's something that we learned about, we didn't know about that until it was brought up. We've learned about it and we have been talking with the school system about how do we change this? How do we stop announcing that students' counselors are in the office to see them? You know, what do we do differently? So, that's just one example.

Although quality assurance is an important role of program evaluation, it is often difficult for nonprofit staff to implement. Jessie Franklin, a senior executive with a provider nonprofit, reporting receiving feedback from peers regarding their need to understand, implement, and use program evaluation for quality assurance purposes. Jessie noted that the discussion of program evaluation efforts with colleagues focused on the struggle with quality assurance issues related to goal attainment and the overall efficacy of their programs:

> I would say it comes up fairly frequently, once again, we don't necessarily get into. . . their specific programs, but generally, "how do I know if we're meeting our goals? How do I help my staff in meeting my goals?" If I'm writing a grant, "how do I know what expectations?" that type of thing.

Leaders also recognized that data from program evaluation efforts inform allocation of resources. For providers, the information helps managers determine if programs should continue or be eliminated. However, most of the discussion regarding resource allocation centered on the use of program evaluation information by funders to continue financial support or not. Morgan Evans, a senior executive with a funding organization, candidly noted this very fact:

> We've got limited resources, and so when you're thinking about, "Where do I allocate my dollars?" You know, that's why the funders said it is so important to know about impact. Because they don't have enough money to say, "all right everybody, do whatever you want to do and let's hope that it's going to work out." I mean, unfortunately, we are in a situation where we've got to pick and choose. So, I think that creates a lot of the fear, which is legit.

Reflective practice. A common theme identified from the analysis of the interviews regarding the role of program evaluation in the Northeast Florida nonprofit sector was the use of reflective practice as a component of organizational learning (the second highest ranked category from the Delphi surveys). Reflective practice was a part of the overall definition for the category. However, reflective practice was the focal point of program evaluation for a majority of the study participants interviewed.

Leaders interviewed noted that the Jessie Ball duPont Fund along with the Community Foundation were the leaders of the reflective practice movement in the Northeast Florida nonprofit sector. The five-year initiative (2000-2005) provided $2,620,933 to 65 nonprofit organizations in the sector. Another 23 nonprofit agencies received ongoing support to continue work produced from their reflective practice projects (Littlepage, KBT & Associates, 2008).

Two of the leaders interviewed provided secondary data related to the reflective practice project. One piece of data was a final report guideline provided to grantees by the funder. The other document was a report produced by the Jessie Ball duPont Fund. The report, "Notes from the Field: Strengthening A Community's Nonprofit Sector" (Littlepage, KBT & Associates, 2008), documented the development in the Northeast Florida nonprofit sector to "nurture a culture of reflective practice" (p. 12). The reflective practice initiative also influenced how the Northeast Florida nonprofit sector utilized program evaluation data. An Executive Director of a funding agency that participated in the project noted, "agencies and funders now use data more to inform the work, rather than anecdotes or assumptions" (Littlepage, KBT & Associates, 2008, p. 17) as a direct result of instituting reflective practice.

Seven years after the reflective practices grants ended, leaders interviewed for this study continued to note the use of reflective practice as a vital component to program evaluation efforts in Northeast Florida. One such leader is Pat Baur, a senior executive with a provider nonprofit. Pat cited reflective practice as the framework used for program evaluation. Pat noted,

> We use a reflective practice model, which actually the Jessie Ball duPont Fund brought to this community . . . and so we use that framework, reflective practice which starts with your evaluation matrix, to then inform, increase our knowledge, "what's working, what's not working?" and to really add it to our knowledge base.

As a proponent of program evaluation and the "power of data," Pat indicated participation in the 2000-2005 reflective practice initiative.

Jordan Duncan, a senior executive with a funding organization, was quick to mention the benefits of reflective practice:

> Well, we really spent some time looking at our processes and we looked at how we exercised our role as a grant maker, and we did a lot of time talking to the grantees …. And in going through that, we realized that we weren't asking the right questions and we weren't learning from the grants, from the body of work that we were administering. So, we really looked at our processes and said, "What needs to change?" So we tried it with [names of organizations], and the other things that we do. So, it just worked really well, the grantees liked it, you know, we listened to them, it changed how our program officers looked at how they evaluate grants. So we just, we kind of applied our own learning to what we were doing, which we think is important.

Chris Carmichael, who also participated in the Delphi surveys, is another senior executive from a funding organization involved with the reflective practice project. Even after the amount of time invested in the Delphi surveys, Chris readily agreed to a follow-up interview. Chris was eager to talk about reflective practice. Chris noted, "It's reflective, it's designed to get you to learn from that particular body of work, that particular investment." Chris further summed up the basic tenets of reflective practice in three simple questions: (a) What do we know? (b) What don't we know? and (c) What are we going to do with this information?

Chris reported being very invested in reflective practice and encouraged feedback from grantees using this framework. However, Chris was quick to acknowledge that other program evaluation efforts are lacking in the Northeast Florida nonprofit sector. Chris noted, "I know we, as a fund, are not participating in that kind of evaluation [sector-wide summative evaluation], and I don't think we as a sector of funders are doing that, on the private side." Chris's observation suggested that although reflective practice is a vital component of program evaluation in the Northeast Florida nonprofit sector, other program evaluation efforts are lacking.

Although not always specifically using the term reflective practice, all of the leaders interviewed talked about the role of program evaluation as an opportunity to promote learning. Taylor Green, a senior executive with a provider nonprofit, explained that being a learning organization is firmly embedded in the culture of the agency where Taylor works. Taylor noted that realigning program evaluation efforts with agency programs (rather than in the fund raising department) changed the emphasis of the role of program evaluation in the organization. Taylor observed, “Because it’s about learning, not about proving.”

Jordan Duncan, a senior executive with a funding organization, viewed learning as a primary role of program evaluation. Jordan reported changing how they manage grants to a focus on learning rather than reporting. Jordan said, “we no longer require them to do reports half way through the grant, we pull them all together and figure out what they are all learning and how can they bounce ideas off of each other… So it’s much more of a conversation than it is an evaluation. It’s working together to say, ‘how can we apply what you’re learning to what you‘re doing?” Jordan also indicated that grantees are encouraged to share what is not working because “we learn more from what we aren’t successful at.”

Although rated high on the Delphi surveys and emphasized in interviews, some leaders were quick to acknowledge that many nonprofits do not use program evaluation information for learning. Jordan observed, “People think when you’re learning you’re not doing, but learning is doing…but some people don’t think, because it is not direct service provision, that you don’t need it, or they don’t value it.” Taylor Green was also alarmed with how many nonprofit provider peers complained about not having the time or desire to do some sort of program evaluation. Taylor’s concern linked directly to organization learning. Taylor commented, “If you are not doing that [program evaluation], how are you learning what’s working and what’s not?” Pat Baur, a senior executive with a provider nonprofit, had similar conclusions:

> I think they [nonprofits] don’t always look at the organizations as learning organizations and the data is just part of that learning. And so they, what I have seen in this community…It’s the summary reports. And no one has really been thoughtful about what it is…It’s just that evaluation is a report to go to a funder. And that’s it. And not as a tool for working smarter.

These perceptive observations seem to indicate that some nonprofit organizations still have yet to move from “a culture of compliance to a culture of inquiry,” as Chris Carmichael, a senior executive with a funding organization, so succinctly described the goal of organizational learning through reflective practice.

Communication with stakeholders. Delphi panelists indicated that a role of program evaluation was a means to communicate with stakeholders, with particular emphasis on funders. Leaders interviewed also brought up communication with stakeholders. Using program evaluation as a tool for communication was particularly important to Morgan Evans. As a senior executive with a funding organization, Morgan was keenly aware of the importance of providing feedback to donors to help increase support. Morgan promoted the use of both quantitative and qualitative data to communicate with stakeholders. Morgan stated a preference for quantitative data:

> As far as communicating with multiple stakeholders, what I continue to learn is that I’m drawn to the numbers, you know, numbers speak to me. And I know the data can be manipulated, I get that, but data, I think, is more validating for me than personal stories.

On the other hand, Morgan recognized the value of qualitative data to communicate to a larger pool of stakeholders. Morgan remarked, “There are a lot of people who really get into the personal stories, so the data is less important, or less impactful to them.” Morgan also reported using statements and examples from the qualitative data to market programs.

Jordan Duncan, a senior executive with a funding organization, also weighed in on the importance of communicating with stakeholders:

> I think, one, the communication piece just can’t be downplayed. You know, if you can’t communicate what it is you're doing to different stakeholders, you really aren’t going to have a program. Communication is a huge piece of this and non-profits often don’t have communications staff, that’s almost seen as a luxury. You know, communications is almost seen as a luxury. So that, it is a really big reason to evaluate.

Jessie Franklin, a senior executive with a provider nonprofit, also noted that using qualitative and quantitative program evaluation data as a mean to communicate to stakeholders and garner financial support was essential:

> So, there was a tremendous amount of support for this idea of "What's the return on my individual investment that I'm making with your organization?" Yet, if you asked people what the most effective way to get people to donate is, it's hearing the stories. So, there's sort of a difference between what they say they want and what they actually end up responding to, so you still need both. Making sure that you have both of those is sometimes tricky . . . how do we tell that story?

Jessie further clarified that part of the story telling process is to communicate the impact the organization has on the community.

In summary, results from the present study found that nonprofit provider and funder leaders perceived the role of program evaluation on a broad spectrum. Results indicated that nonprofits used program evaluation efforts for internal and external purposes. Internal uses for program evaluation efforts included advancing organizational learning, shaping management decisions to include resource allocation, and facilitating quality assurance.

Organizations used information from program evaluation efforts for communication to external stakeholders. The most cited areas of communication included information regarding impact, outcomes, mission validation, organizational credibility, and garnering resources. In particular, as discussed in the next section, a primary role of program evaluation was to enhance the relationship between providers and funders.

Research Question #2

The second research question focused on how the relationships between service providers and funders influence program evaluation efforts in the Northeast Florida nonprofit sector from the perceptions of the leaders in the present study. Results regarding the relationship between providers and funders were mixed. Initial findings for this question came from the first exploratory round of the Delphi surveys. In particular, Question 8 (see Appendix D) asked how funders and providers work together on program evaluation efforts. This was the last question on the survey, and one of the participants (a provider) did not answer the question. Additionally, the question had the fewest responses. A funder Delphi panelist remarked, "not very well." Another funder panelist observed, "We [providers and funders] are not collaborating around evaluation."

Interviews produced more in-depth information. The leaders interviewed had differing opinions regarding the relationships between providers and funders. An analysis of the interview data found that funders focused on building partnerships and collaboration. On the other hand, providers focused on the funder requirements. Ultimately, group affiliation (providers or funders) appeared to influence the perception of study participants regarding the relationships between providers and funders and may explain some of the diverse feedback on the subject. Moreover, providers and funders are not a homogenous group and provided responses to specific organizations and entities (e.g., private versus public funding).

Unequal partners. One does not have to dig deep to see the differences between funder and provider organizations. An observation of office location for each group is just the beginning in understanding the differences in emblems of status and access to resources. All of the funders interviewed had offices in prime realty spots in Northeast Florida. On the other hand, most provider organizations were located in neighborhoods with a diversity of economic classes. Few had scenic views.

The unequal distribution of power between providers and funders is not lost on Chris Carmichael, a senior executive with a funding organization. Chris noted that many nonprofits have a negative view regarding funders because of the power differential. Chris observed, “I think what nonprofits say about the foundation community, funder arrogance. I think it’s real. I think it’s true. We own all the power. Let’s be honest.” Jordan Duncan, a senior executive with a funding organization, also noted the negative terms often used to describe some funders. Jordan remarked, “I think there’s this term, ‘arro-grant making,’ you know, grant makers are being arrogant.”

In contrast to some negative views about funders, Jordan and Chris described how their organizations made it a priority to minimize the power differential and develop strong partnerships with providers. Chris described the type of attitude funders need in order to build positive alliances:

> I think you have to have a natural humility, a natural interest in your peers being as successful as you are, that you cannot see this as a competition…And, can, you know, park their own arrogance, and be generous in spirit, and, you know, understand that this is a human business and that there is a lot of give and take and there are a lot of nuances to this work.

Jordan also reported making an effort to develop partnerships with grantees. Jordan even demonstrated this by intertwining fingers when speaking about the subject:

> It’s very much a partnership because philanthropy is only successful if the nonprofits serving the individuals are successful. … So we’ve definitely seen our role as a partnership, and if the organizations aren’t successful, then we aren’t successful, then our donors don’t feel that they're successful. So, it’s really important that you look at it as a partnership. It’s not a power relationship.

Pat Baur, a senior executive with a provider nonprofit, noted that communication between providers and funders is the key to equalizing the relationship and moving forward:

> I think part of it is having some dialogues with the major funders in our community and the nonprofit executives, to have that kind of open dialogue, to say, “We want to be able to come and have these conversations and be given the opportunity to be heard,” and I think it has to be both ways. I think we have to have that conversation with the funding community, so that it doesn’t seem like such an uneven distribution of power.

Interestingly, while Pat described "the uneven distribution of power" between the two groups, Pat's hands were pointed at each other in a parallel fashion with one higher than the other. However, Pat's hands moved on an equal plane when describing personal communication with funders. At the time, I pointed this out to Pat, who had not realized the change in the position of the hands. Pat responded,

> Did I? Well, I feel that I'm equal to them….I think it is a privilege to be part of the organization and the work we do. If funders don't see that, they're probably not going to be a good partner with us.

Yet, in spite of Pat, Jordan, and Chris's optimistic experiences of partnership between the groups, those interviewed acknowledged that funders retain the power in the relationship, which makes authentic partnership more difficult to achieve.

Funder requirements. Both funders and providers observed that program evaluation efforts primarily exist in Northeast Florida because of requirements from funders. Morgan Evans, a senior executive with a funding organization, remarked, "It's the funders that are requiring the program evaluation…. I think it really is being pushed from one direction, that's my experience." Morgan even reported the belief that most provider organizations would not engage in a reflective assessment of their services if not prompted by the funding community.

Jordan Duncan, a senior executive with a funding organization, agreed that funders require program evaluation and described how the demands from government funders were particularly burdensome to providers:

> Now certainly public funders, so government funders, there are a lot of requirements. Some of them I think are very unrealistic or difficult for nonprofits. They require a lot of evaluation but they don't really fund you to have positions to be able to do that. They don't understand the cost to an organization to be able to measure anything. You know, every time you add a measurement you add cost. And I think that a lot of the government funders do not necessarily compensate nonprofits, especially when they refuse to fund administration and overhead, so how do you then do that kind of work? Because those people typically aren't the people that are on the ground delivering the service. So it

> makes it difficult for nonprofits, but I think that it's here to stay. I don't think that it's going to change. I think it's probably going to increase.

Jessie Franklin, a senior executive with a provider nonprofit, characterized the relationship between government funders and provider as being more adversarial.

> I think probably our biggest struggle is government funders and making them understand what evaluation is, what it isn't, what it can help them with, and what it can't help them with. I think we continue to struggle with that, especially on the local level. On the federal level, they, I think, tend to back off a little bit, I think they recognize that there's some limitations to what you can find out through evaluation. At the local level, they seem really determined to catch us doing bad things. You know, to figure out a way that we're scamming them out of their money. As opposed to viewing their investment as a vital part of supporting the work that needs to happen in this community.

Study participants also included private funders (e.g., foundations, corporations) as responsible for requiring program evaluations from providers. Pat Baur, a senior executive with a provider nonprofit, noted the following:

> I think that the sector, and I think it's the nonprofit sector in general, I think that many people see research as an afterthought or "I have to do it because this grant says I have to do it," or "because my funding source is making me do it." So it gets done after the fact as opposed to it being integrated into the day-to-day work of the organization.

Taylor Green, a senior executive with a provider nonprofit, perceived a possible danger in a funder-directed partnership:

> Well, certainly there is the potential for there to be a negative impact in that you may be tailoring services to specifically what a funder needs. If you're chasing the money, that may or may not be a good thing…. I mean, are you really doing things outside of your strategic plan? Are you pushing that in order to keep money flowing? That could be tough. …So it's not that you can't sometimes change what you're doing in order to meet the needs of a funder, it's when you really get too far afield, I think that you always have to wonder, worry about.

Results from Delphi panelists also indicated how funders often drive program evaluation efforts through service delivery expectations and reporting requirements. Most significantly, panelists noted the lack of consistency of funders' expectations of provider organizations. Providers are required to report on multiple metrics to multiple funders. In the Delphi surveys, an executive director of a large provider nonprofit noted, “Each funder uses different tools for program evaluation.” Another Delphi panelist from a provider nonprofit noted, “There is a challenge for funders to devise a way for all of the organizations that they serve to provide information in a uniform way that will allow the programs to be compared to one another.” Yet another Delphi funder panelist had a similar observation, “Everyone has different reporting requirements. I imagine it’s very time consuming for providers.” Jessie Franklin, a senior executive with a provider nonprofit, noted the following negative impact this may have on service delivery:

> And therein lies the dilemma, especially when you get to nonprofits having to respond to different funders, different expectations, what evaluation looks like and then they're spending a lot of time trying to come up with those rather than working on the programs that they're supposed to be running.

Delphi provider panelists noted a particular funding organization that imposed evaluation metrics on providers. Taylor Green, a senior executive with a provider nonprofit, also noted this funding organization’s involvement and requirements regarding program evaluation efforts:

> [name omitted for confidentiality] has a significant impact on our evaluation processes and prescribes many of the outcomes we measure. They are keenly interested in the processes we use, the appropriateness of sample size, etc. Since they ask many agencies to achieve the same outcomes with the same participants, one might wonder why they don’t [do] some of the data collection?

Collaboration and communication. A predominant theme from the results of the Delphi surveys and the interviews was the different ways funders and providers worked together regarding program evaluation efforts. Some Delphi panelists noted that both groups cooperated to share resources such as common performance measures, program evaluation strategies, data access solutions, training, and results. Study participants (Delphi panelists and interviewees) noted that providers and funders preferred program evaluation methods developed by and agreed upon by both groups. However, one Delphi provider panelist noted that cooperation among providers regarding data collection was limited due to confidentiality concerns.

Several of the Delphi responses and interviewed leaders noted specific organizations that contributed to these collaborative efforts. These included the Children's Commission, Donors Forum of Northeast Florida, Florida Philanthropic Network, Emergency Services Homeless Coalition, the Northeast Florida Nonprofit Center, and United Way of Northeast Florida. In particular, study participants gave credit to the Community Foundation and the Jessie Ball duPont Fund for their efforts in building collaboration between funders and providers in the Northeast Florida nonprofit sector. Jessie Franklin, a senior executive with a provider nonprofit, noted the high level of cooperation in the Northeast Florida nonprofit sector:

> I actually think that we have a really positive working relationship between our local funding community and our local nonprofits, for the most part. I personally happen to think that it's driven by the duPont Fund and the Community Foundation. The leadership of those organizations has really influenced how most funders in the community relate to their [provider] organizations.

An analysis of the data from the interviews indicated a recurring theme of trust between providers and funders as a key element of their relationship that has positive outcomes for program evaluation strategies. Central to building trust between providers and funders was on-going communication. Jordan Duncan, a senior executive with a funding organization, described the need for reciprocal communication between providers and funders:

> I think having an honest conversation, the nonprofits pushing back and asking the right questions, the funders pushing back and asking the right questions, and being able to have that dialogue. I think it's important. …. And having it be an ongoing conversation rather than an episode that happens at the end of the grant.

However, Jordan noted that these discussions were not easy at first because providers were skeptical about being honest about the results of their work. As Jordan note, "We were standoffish." In response, Jordan's organization made a concerted effort to engage providers in more discussions without being punitive (e.g., removing funding) when providers failed to meet expected results:

> Then word starts to travel that we really meant what we said and we learned some things along the way…. It really did help us to be, I think, much better at what we do, and more supportive of the sector, you know.

Jessie Franklin and Pat Baur, senior executives with provider nonprofits, had similar conclusions regarding communication as critical to building trust between providers and funders. Jessie noted that "our local funding community is pretty supportive, willing to have dialogue with [provider] organizations, and looking for ways to kind of strengthen the sector rather than put additional burdens on them." Similarly, Pat remarked, "It's somehow about having the conversation with the funding community where, you know, I've been fortunate because I've had the conversation with the funders to say, 'you know, we didn't hit the mark on this, but here's what we learned.'" Even in situations where funding was removed, Morgan Evans, a senior executive with a funding organization, noted, "We have been forthright in communicating why the decision was made, and people may disagree with the decision that's made but they can't say, 'all right, you didn't completely communicate this to me.' They might just have a different opinion about it."

Program evaluation culture. An analysis of the data also found that the relationship between providers and funders produced an overarching culture related to program evaluation efforts. Primarily, the funder-driven aspect of program evaluation seems to put the focus on compliance to funder reports with the emphasis on outputs, grant requirements, and (ultimately) ensuring continued funding. Pat Baur, a senior executive with a provider nonprofit, viewed this as a major concern for the nonprofit sector in Northeast Florida, which Pat believes can lead to mission drift:

> They are doing the work to keep their funding streams going, as opposed to me, where I see the work; it's about moving your mission forward. And using evaluation as a way of propelling that mission forward…. This is where we get into, "mission follows funding," as opposed to "funding follows mission," this to me, says it. And that's the culture of our, of Northeast Florida…. I think it's what's creating the culture, that missions of these organizations, of organizations and vision, gets skewed because they are following the funders. As opposed to the funding should be trying to move missions forward.

Morgan Evans, a senior executive with a funding organization, also worries that this focus of program evaluation on outputs may also lead to unintended consequences for program participants. Morgan recalled a local meeting where the Mayor of Jacksonville was talking to the nonprofit sector and shared the following example:

> He [Mayor of Jacksonville] said, "I'm out in the communities. I am talking with people all the time…. They do not trust organizations and they feel like they are often a statistic; they are a number for funding." And that really gave me pause, because I wonder how often we are either guilty of doing that or guilty of creating the culture within which that occurs. You know, we've got these outputs that you need to have and, is that undermining the ability to treat people like people?

Study participants also noted how the emphasis on metrics has also created a culture of fear regarding program evaluation efforts. Morgan Evans noted, "We have some organizational cultures that are just afraid of evaluation themselves." The fear is focused around losing funding and/or "people will find out that we are not doing well" as Pat Baur observed. To take this a step further, Morgan Evans observed that, as a community, we use program evaluation against each other:

> I think it goes back to, we live in a society where we blame, we love to blame. And, you know, holding people accountable is, you know, "it's your fault, you're out," which isn't what it needs to be. I was in a meeting not too long ago where someone said, "We're exploiting each other's failure." And we've got to stop doing that. We have got to start realizing that what you contribute is important, what I contribute is important, and how do we work together, you know, to make our contributions even better.

On the other hand, study participants did mention positive aspects of culture in the Northeast Florida nonprofit sector. First, the reflective practice project initiated by the Jessie

Ball duPont Fund and the Community Foundation fostered a "culture of inquiry." The reflective practice project continues to evolve through a sector initiative, Moving Forward Together, facilitated by the Nonprofit Center of Northeast Florida:

> Moving Forward Together is a sector-owned and -driven initiative that connects and strengthens local nonprofits. Spearheaded by the Nonprofit Center and funded by the Community Foundation in Jacksonville and the Jessie Ball duPont Fund, Moving Forward Together spurs collective, innovative, and sustainable action to achieve progressive and lasting social change. (Nonprofit Center of Northeast Florida, 2013b)

One of the focuses of Moving Forward Together is creating positive relationship within the sector. A second focus is measuring the value and impact of nonprofits. Leaders interviewed acknowledged that the Nonprofit Center of Northeast Florida and the Moving Forward Together project are vital to continuing the dialog regarding program evaluation efforts as well as enhancing an overall positive nonprofit sector-wide culture.

Jessie Franklin, a senior executive with a provider nonprofit, observed nonprofit providers may have initially viewed funder-directed program evaluation requirements as negative. Jessie noted, however, that the culture evolved to accepting and understanding the need for program evaluation:

> I think it mostly started being funder driven, but it started with, there was a lot of movement in the philanthropy world, then they said, "Well, how do we know we are making a difference with all this money we are investing? So let's start asking questions." And they started putting pressure on non-profits to begin with, and then as they [nonprofits] got more comfortable with it and used to it, they started embracing on their own. And now, boards are asking those questions, or organizations are asking themselves those questions . . . . They're constantly going, "Is this really telling the story that I want it to tell? Do I really know if I am making a difference? How's everyone else doing it? How can we figure out better ways to do it?"

Other leaders noted how funding organizations have created opportunities for dialog around deliverables and provided training opportunities. These efforts helped move the culture in a positive direction regarding program evaluation.

From the results of the present study, it is obvious that the relationship between providers and funders influences the role of program evaluation efforts. From specific funders' requirements to creating a culture receptive to program evaluation, the influence of relationship between the groups was evident. Additionally, as is discussed in the next section, the results of the present study found that the relationship between providers and funders had significant bearing on program evaluation capacity.

Research Question #3

The third and final research question for the present study focused on the perceptions of nonprofit and funding organizations' leaders regarding the development of program evaluation capacity in the Northeast Florida nonprofit sector. To answer the question, I examined the development of capacity for program evaluation from three different angles. The first approach was to ascertain from Delphi panelists the essential capacity elements needed for program evaluation. The second approach examined to what degree the capacity elements from the Delphi survey existed in the Northeast Florida nonprofit sector. Finally, interviews provided rich data regarding how to build program evaluation capacity.

Essential capacity elements. From my analysis of the first Delphi survey, I identified eight overarching essential capacity elements from the over 100 responses from the capacity related questions. Using a five-point rating scale from 1 (*strongly disagree*) to 5 (*strongly agree*), panelists rated each category. Panelists provided ratings for the categories during the second Delphi round, fourth Delphi round, and interviews phases of the study (see Table 11).

Table 11

Program Evaluation Essential Capacity Elements Rating Responses Frequencies

Categories	Rating Responses				
	1	***2***	***3***	***4***	***5***
Sufficient Time	0	0	0	3	9
Sufficient Financial Resources	0	0	0	5	7
A Positive Culture	0	0	0	6	6
Functional Program Evaluation Designs/Methods	0	1	0	3	8
Sufficient Human Resources	0	0	3	2	7
Realistic Expectations from the Philanthropic Community	1	0	1	4	6
Ongoing Collaboration	0	0	3	5	4
Ongoing Training	0	0	5	3	4

Note. Anchors for the scale were 1= *strongly agree*; 2 = *disagree*; 3 = *neither agree or disagree*; 4 = *agree*; 5 = *strongly agree*.

I ranked the findings according to the means of each category from all of the panelists' (providers' and funders') ratings. As noted in Table 12, seven of the eight categories had mean ratings higher than 4. The essential capacity element category with the highest mean was "sufficient time." This particular category also had least variation with a standard deviation of .45. The lowest rated category (ongoing training) had an overall mean of 3.92. However, the category with the greatest variation with a standard deviation of 1.19 was "realistic expectations from funders." The difference between the highest and lowest ranked categories was less than one (.83).

Table 12

Delphi Survey Results of Essential Program Evaluation Capacity Elements

Categories	Combined *M* (*SD*)	Funders *M* (*SD*)	Providers *M* (*SD*)
Sufficient Time	4.75 (.45)	5.00 (0.00)	4.57 (.53)
Sufficient Financial Resources	4.58 (.51)	4.60 (.55)	4.57 (.53)
A Positive Culture	4.50 (.52)	4.60 (.55)	4.43 (.53)
Functional Program Evaluation Designs/Methods	4.50 (.90)	4.60 (.55)	4.43 (1.13)
Sufficient Human Resources	4.33 (.89)	4.60 (.89)	4.14 (.90)
Realistic Expectations from the Philanthropic Community	4.17 (1.19)	4.40 (.89)	4.00 (1.41)
Ongoing Collaboration	4.08 (.79)	4.40 (.55)	3.86 (.90)
Ongoing Training.	3.92 (.90)	4.40 (.89)	3.57 (.79)

Note. In those instances where the mean is the same but the standard deviations are different, the category with the lower standard deviation is ranked higher.

The ranking per subgroup of providers and funders indicated that the need for sufficient time had the highest overall mean for both providers and funders. However, the need for essential financial resources had the same overall mean and standard deviation for the providers' subgroup. Although not in the exact ranked order, both subgroups had the same categories ranked as the top three essential capacity elements needed. Similarly, both subgroups had the need for ongoing training ranked last (see Table 12 for more detailed information).

I specifically interviewed two Delphi panelists because of their outlier ratings related to program evaluation capacity. At the onset of each interview, I explained that I wanted to conduct a follow-up interview to learn more about their individual responses and to make sure they understood the intent behind the essential capacity elements section of the survey. I explained that, for the second and fourth Delphi rounds, the capacity elements section required panelists to rate each category based on a determination if each element was needed for program evaluation efforts whether the capacity element existed or not. I learned that one panelist misunderstood the intent of the ratings. The Delphi panelist thought rating essential capacity elements was in relation to the level each element existed in the Northeast Florida nonprofit sector. The Delphi panelist wanted to change the ratings upon realizing the misunderstanding. The other outlier Delphi panelist, a senior manager of a provider organization, affirmed understanding the intent of the essential capacity elements section. However, this Delphi panelist also wanted to change a few final ratings after further consideration.

I allowed changes from both of these for a few reasons. First, as the focus of the present study was to learn the perceptions from the participants, I wanted to make sure the ratings represented their perceptions. Second, the present study is purposefully a qualitative design with participants being the drivers of the findings. Consequently, I honored their requests to change their ratings. Finally, I designed the interview phase of the study to follow the Delphi surveys phase in the event of outlier responses that needed explanation or adjustment. Combining the results from the Delphi on-line surveys with the follow-up interview ratings of Delphi surveys outlier responses helped strengthen the present study. B. Johnson and Christensen (2008) described this as a mixed method aspect of research that reduces gaps in research design.

Current program evaluation capacity. As noted in the previous section, it was important to understand the needed program evaluation capacity elements from the perceptions of providers and funders. However, just as important was exploring the degree to which the agreed upon essential capacity elements were present in the Northeast Florida nonprofit sector. Therefore, the third Delphi round provided panelists with the opportunity to rate to what degree program evaluation capacity elements are currently present in the Northeast Florida nonprofit sector. Essential capacity elements included were those categories that had an overall mean rating of 4 or higher from the second Delphi round (see Appendix H). Those categories included (a) sufficient time, (b) functional program evaluation designs/methods, (c) sufficient human resources, (d) a positive culture, (e) sufficient financial resources, and (f) ongoing collaboration. At the time, the “realistic expectations from the philanthropic community” category had an overall mean of 3.92 and the “ongoing training” category mean was 3.58. Due to not reaching an overall mean of 4 or higher (*agree*, *strongly agree*), I chose not to include them on the ratings for the third round Delphi. In hindsight, because the overall means changed after final ratings from on-line surveys and interviews, I should have included all of the essential capacity elements for the third and fourth Delphi survey rounds.

For this round, I asked Delphi panelists to rate the degree to which the elements were currently present in the Northeast Florida nonprofit sector (see Appendix H). The rating scale ranged from 1 (*usually not present*) to 5 (*generally present*). After panelists completed the third round, I provided each of them (for the fourth Delphi round) their individual responses compared to the means of the sub-groups (providers and funders) and combined group. I also provided comments from other Delphi panelists. At this point, panelists could change and/or provide comments related to their individual ratings (see Table 13).

The final results for the current level of program evaluation capacity in the Northeast Florida nonprofit sector revealed that only one capacity element (positive culture) had a combined (providers and funders) mean higher than 3. On the other end of the continuum, financial resources had the lowest combined (providers and funders) mean of 2.17. The other capacity elements also had means less than 3. Overall, the results seemed to indicate that leaders of human services provider and funding organizations perceived that capacity for program

Table 13

Current Program Evaluation Capacity Rating Frequencies

Categories	***Rating Responses*** 1	2	3	4	5
A Positive Culture	0	1	8	2	1
Functional Program Evaluation Designs/Methods	0	2	10	0	0
Ongoing Collaboration	1	1	10	0	0
Sufficient Time	1	3	6	2	0
Sufficient Human Resources	2	5	4	0	1
Sufficient Financial Resources	3	5	3	1	0

Note. The anchors for the scale were 1 = usually not present; 2 = often not present; 3 = *sometimes present*; 4 = *often present*; 5 = *generally present*

evaluation efforts in Northeast Florida was lacking (see Table 14).

Both funders and providers had a positive culture as the highest overall mean rating. Similarly, the both groups' category with the lowest overall mean rating indicated a lack of financial resources. It is interesting to note the funders overall means for the current levels of capacity for each category were lower than providers overall means (see Table 14 for detailed information).

Program evaluation capacity development. The leaders interviewed for the present study agreed with the Delphi survey results regarding the needed essential capacity elements and the current level of program evaluation capacity. As a follow-up, one of the questions put forward during the interview phase of the study asked the leaders to consider strategies for

Table 14

Delphi Results of Current Program Evaluation Capacity

Categories	**Combined** ***M*** **(*SD*)**	**Funders** ***M*** **(*SD*)**	**Providers** ***M*** **(*SD*)**
A Positive Culture	3.25 (.75)	3.20 (.45)	3.29 (.95)
Functional Program Evaluation Designs/Methods	2.83 (.39)	2.80 (.45)	2.86 (.38)
Ongoing Collaboration	2.75 (.62)	2.60 (.89)	2.86 (.38)
Sufficient Time	2.75 (.87)	2.20 (.84)	3.14 (.69)
Sufficient Human Resources	2.42 (1.08)	2.00 (.71)	2.71 (1.25)
Sufficient Financial Resources	2.17 (.94)	1.60 (.55)	2.57 (.98)

Note. In those instances where the mean is the same but the standard deviations are different, the category with the lower standard deviation is ranked higher.

developing program evaluation capacity. An analysis of the data suggested congruency among the leaders interviewed regarding specific measures that could further develop program evaluation capacity in the Northeast Florida nonprofit sector. The key ideas from the present study included the following steps for increasing program evaluation capacity in the Northeast Florida nonprofit sector:

- Identify a person or organization to “champion” program evaluation efforts.
- Educate both providers and funders regarding the value of program evaluation and the associated costs.
- Gather information and study other nonprofit sectors/organizations that promote successful program evaluation strategies.
- Build on collaborative efforts with colleges/universities, providers, and funders.

Leaders interviewed acknowledged that any intentional capacity building endeavor regarding program evaluation would be most successful if there were a person and/or an organization willing to champion the effort. The champion would be someone who could garner support and resources. Chris Carmichael, a senior executive with a funding organization, noted the type of person needed, “Someone who will stay in for the long haul, because it takes forever. And I think has to bring some capital to it. They can’t just bring the brain power.” Jessie Franklin, a senior executive with a provider nonprofit, also described the type of person needed within organizations to create momentum for program evaluation:

> My personal opinion on it is that if you're looking for someone who has expertise and is passionate about evaluation, you're probably looking at a new person for an organization, which means finding the funding to expand your staff to actually include that expertise.

Although Jessie focused on organizations hiring staff members to champion program evaluation, others suggested that the Nonprofit Center of Northeast Florida would be a logical choice to begin the discussion regarding program evaluation in the sector.

Leaders interviewed determined that educating providers and funders regarding the many facets of program evaluation was another key element to building capacity. Part of this education included showing how program evaluation information can positively influence effectiveness, efficiency, and increase benefits to constituents. According to those interviewed, this was a key concern for the funding community. Jordan Duncan, a senior executive with a funding organization, also noted that educating the funding community was critical. Jordan explained, “I think being clear with funders [about] what the benefits are and what the drawbacks are. Then helping the sector to better communicate what it is that they need.” Jessie Franklin, a senior executive with a provider nonprofit, also noted the need for educating funders:

> So, maybe a little bit of educating funders about the importance of funding something like evaluation . . . and then giving non-profits the resources and the permission to kind of use that as a way to expand their work. And probably linking it back to how evaluating actually ends up giving you the ability to serve more people or to meet your mission more effectively, rather than taking dollars away from running the program.

However, Chris Carmichael, a senior executive with a funding organization, had some hesitation regarding whether educating funders would help increase program evaluation funding. Chris remarked, "I just think we could have a great chat about it and why it would be helpful to us but, at the end of the day, I don't think we'd fund it."

Another suggestion for capacity building focused on learning from other nonprofit sectors that seem to have a higher degree of cohesiveness and capacity regarding program evaluation. Pat Baur, a senior executive with a provider nonprofit, noted that other cities seem to have developed consistency around program evaluation strategies, which is reflected in their RFPs (request for proposals). Jordan Duncan noted the Foundation Strategy Group and GrantMakers For Effective Organizations as two groups that provide support and information regarding program evaluation strategies.

The final important element to developing program evaluation capacity is leveraging the collaborative culture that exists in Northeast Florida. As noted previously, the five-year reflective practice movement spearheaded by the Jessie Ball duPont Fund and the Community Foundation created a strong collaborative base in the sector. In fact, a study regarding this effort reported, "The work has strengthened relationships within the nonprofit sector and between donors" (Littlepage, KBT & Associates, 2008, p. 17).

Even years later, the positive results of the project can be seen through the creation of the Nonprofit Center of Northeast Florida and the positive comments from leaders regarding the level of cooperation in the sector. Chris Carmichael, a senior executive with a funding organization, is one example of a leader who noted the positive collaboration culture. Chris noted, "I think we built something really important here and very special." Even so, the leaders interviewed noted that building continued collaboration around program evaluation was necessary to increase capacity.

Leaders interviewed were of the opinion that creating a stronger collaboration with local colleges and universities was essential to developing capacity. Pat Baur, a senior executive with a provider nonprofit, was surprised at the limited interaction between the Northeast Florida nonprofit sector and universities:

> The disconnect between the nonprofit community and the universities in Jacksonville was a little startling to me, because it was very different when I was in [omitted] and it was very different when I was in [omitted] and [omitted]. There was just an alignment with the nonprofit community and the universities who had this expertise of evaluation, and here it just wasn't . . . . I think there's been a little bit of progress, maybe, made in bringing those two entities together, but I think that is really missing in this community.

Leaders agreed that an increased engagement of nonprofits with the academic community should not be limited to hiring them solely for conducting evaluations. Only using universities to conduct evaluations did not address other capacity issues. More importantly, the findings suggested that nonprofit leaders should expand beyond the use of colleges and/or universities for conducting evaluations to include (a) provide training, (b) create program evaluation designs, (c) serve as technical assistants, (d) provide interns, (e) and provide program evaluation classes specifically tailored to the nonprofit sector.

Continued collaboration between the provider and funding communities was also another vital element to capacity building. First, both groups should continue to make time to discuss evaluation priorities. As Morgan Evans, a senior executive with a funding organization, noted, "It's incumbent upon funders in the community and the nonprofits to work together and figure out how that time can be made available." Study participants cited the United Way of Northeast Florida as a prime example of advancing these types of conversations. For example, United Way brought together different agencies to determine common metrics for measurement for children at risk. Furthermore, staff of the United Way seek out feedback from providers to assist with interpreting data.

One of the prevailing issues around collaboration and capacity development focused on increasing funding for program evaluation. However, study participants did not feel pouring money into program evaluation efforts was the first step to sustaining capacity. Indeed, they firmly believed that increasing collaboration between providers and funders was the bedrock for generating additional resources. Jordan Duncan, a senior executive with a funding organization, noted that collaborative efforts could expand program evaluation capacity:

> If you did more collaboration about this, between the providers and the philanthropic sector, you could address some of these things like sufficient resources, sufficient time, human resources, you know. If you had more frank conversations around it and understood where each other was on it, you might be able to address some of these.

It seemed that Jordan and the other leaders interviewed understood how continued engagement between providers and funders directly affects capacity-building efforts.

In summary, the development of program evaluation capacity is multi-faceted. First, it is important to know what factors contribute to capacity. The results of the present study indicated eight essential elements needed for program evaluation. Second, an assessment of the current level of each capacity components is necessary to understand where to target capacity building efforts. Results from the present study found that current program evaluation capacity was deficient. Finally, the development of program evaluation capacity is intentional work. Results from the present study indicated several specific steps that could help develop program evaluation capacity in the Northeast Florida nonprofit sector.

Summary

The present study explored the perceptions of human services provider nonprofits and funding organizations regarding the role of and capacity for program evaluation in the Northeast

Florida nonprofit sector. Findings from the present study provided information in three main areas: (a) the different roles of program evaluation, (b) the impact of the relationship between providers and funders on program evaluation efforts, and (c) the development of program evaluation capacity.

The present study consisted of two phases. The first phase of the study was a multi-round Delphi surveys conducted with leaders from both the provider and funding communities. Results from the Delphi surveys indicated eight unique categories related to the role of program evaluation: (a) assess impact of program, (b) advance organizational learning, (c) cultivate funding collaborations, (d) inform program management decisions, (e) enhance communication with multiple stakeholders, (f) facilitate quality assurance, (g) determine resource allocation, and (h) validate organizational credibility.

Findings from the Delphi surveys also provided eight essential capacity elements needed for program evaluation efforts. The essential program evaluation capacity elements were (a) sufficient time, (b) sufficient financial resources, (c) a positive culture, (d) functional program evaluation designs/methods, (e) sufficient human resources, (f) realistic expectations from funders, (g) ongoing collaboration, and (h) ongoing training. Of these categories, except for training and realistic expectations, panelists perceived the current level of program evaluation capacity in the Northeast Florida nonprofit sector as lacking.

Findings from the interview phase of the study mirrored the Delphi results while providing robust narrative regarding the state of the sector relative to program evaluation. Additionally, data from the leaders interviewed revealed valuable insight regarding the level of collaboration between providers and funders.

Chapter 5 presents a summary of the results along with a discussion of the findings. The discussion includes major conclusions and limitations of the present study. Implications and recommendations for further research and practice are also presented.

Chapter 5: Conclusion

The purpose of the present study was to explore the perceptions of leaders from human services nonprofits and funding organizations regarding the role of and capacity for program evaluation in the Northeast Florida nonprofit sector. In this final chapter, I provide a summary and discussion of the results. I then offer primary conclusions and recommendations. Finally, I end the chapter and the dissertation with concluding comments.

Summary of Results

The results of this two-phase study provided insight into the perceptions of leaders in the Northeast Florida nonprofit sector regarding the role of and capacity for program evaluation. First, participants from the Delphi surveys phase of the study clearly indicated eight different yet related roles for program evaluation: (a) assess impact of program, (b) advance organizational learning, (c) cultivate funding collaborations, (d) inform program management decisions, (e) enhance communication with multiple stakeholders, (f) facilitate quality assurance, (g) determine resource allocation, and (h) validate organizational credibility. Responses to seven of the eight categories had generally high scores. Responses for the remaining category (validate organizational credibility) had the lowest score. Although these eight categories are distinct, an analysis of the interviews revealed a synthesis of ideas regarding the role of program evaluation into four broad categories: (a) impact and mission validation, (b) decision-making tool, (c) reflective practice, and (d) communication with stakeholders.

The second area of results related to the capacity for program evaluation. The Delph

survey results produced eight elements of capacity essential for program evaluation efforts: (a) sufficient time, (b) sufficient financial resources, (c) a positive culture, (d) functional program evaluation designs/methods, (e) sufficient human resources, (f) realistic expectations from the philanthropic community, (g) ongoing collaboration, and (h) ongoing training.
Responses to seven of the eight essential capacity elements had generally high scores. Responses to the remaining element (ongoing training) had the lowest score.

A further examination of the capacity for program evaluation concentrated on the current level of program evaluation capacity in the Northeast Florida sector and the steps needed to develop capacity. The results from the Delphi surveys indicated that program evaluation capacity in Northeast Florida was lacking. The category with the highest score was a positive culture. The category with the lowest overall score was sufficient financial resources.

Results from the interviews provided the steps necessary for developing program evaluation capacity. The first step was to identify a person or organization to champion program evaluation efforts. The second step centered on educating the nonprofit sector on the value and associated costs of program evaluation. The third step was to study information and model nonprofit sectors/organizations that demonstrate positive program evaluation strategies. Finally, leaders determined that building collaborations within the nonprofit sector and with higher education institutions was necessary for developing program evaluation capacity.

Discussion of Results

The results from the present study parallel a review of the literature regarding the role of and capacity for program evaluation in the nonprofit sector. However, many elements of the present study suggested findings and strategies unique to the Northeast Florida nonprofit sector. I organized the format for the discussion of the results of the present study by research question.

The Role of Program Evaluation

As mentioned in Chapter 4, the results for the role of program evaluation included the definition, different approaches, and the different uses for program evaluation. Not surprisingly, answers from Delphi panelists regarding the different program evaluation approaches used a mixture of program evaluation designs (e.g., longitudinal studies, quasi-experimental, empowerment models) and methods for collecting data (e.g., interviews, focus groups, surveys). The different answers from the present study regarding the definition and approaches of program evaluation are similar to other studies. Earlier studies (Carman, 2007; Carman & Frederick, 2008) found that nonprofit leaders used a variety of approaches and methods that they considered evaluations. The approaches ranged from financial audits to programmatic evaluations, with most efforts focused on creating reports for boards. From this research, the authors concluded that the broad interpretation of approaches indicated a lack of knowledge among nonprofit leaders regarding program evaluation strategies. Similar findings came from the present study as well.

However, participants in the present study did not include any financial audits or activities as part of program evaluation. This finding indicated that the leaders involved in the present study understand the differences between financial audits and program evaluation efforts. Furthermore, the focus of the majority of the study participants was on internal program evaluation efforts rather than outside regulators. The difference with the present study and others (Carman, 2007; Carman & Frederick, 2008) is the use of the open-ended questions from the first Delphi survey round. The design of the present study allowed panelists to create their own categories regarding the role of program evaluation rather than using predetermined evaluation categories or activities. Consequently, participants in the present study did not seem to regard financial activities or audits as types of program evaluation. This was a result clearly different from the other research mentioned.

The eight categories related to the role of program evaluation identified from an analysis of the data mirrored other studies (Carman & Fredericks, 2008; Eckerd & Moulton, 2011; Lee et al., 2008). The similar roles for program evaluation from the present study compared to the others were the following: (a) assessment of impact and outcomes, (b) a tool for programmatic and management decisions, (c) quality assurance/program improvement, (d) communication with stakeholders, (e) strategic planning, (f) a way to garner financial resources, and (g) meeting funding requirements .

The results from the Delphi surveys from the present study regarding the highest and lowest rated role of program evaluation categories were similar to results from a study by Eckerd and Moulton (2011). The authors found that 71% of nonprofits surveyed used program evaluation to assess the impact of program, which correlates to the category with the highest mean in the present study. Additionally, Eckerd and Moulton (2011) found that using program evaluation to determine mission alignment was not a priority among nonprofit organizations. In the present study, the category that included assuring mission alignment had the lowest overall mean. The results of the present study regarding the different roles of program evaluation also parallel factors associated with organizational accountability in nonprofit management literature. For example, Ebrahim (2010) noted “five broad . . . accountability mechanisms used by nonprofits in practice: reports and disclosure statement, evaluations and performance assessments, industry self-regulation, participation, and adaptive learning” (p. 107). Notably, several of the roles of program evaluation resulting from the present study correspond to these elements of accountability. In particular, the following roles of program evaluation can be tools (as described by Ebrahim) for providing information for accountability efforts: (a) assess impact of program, (b) facilitate quality assurance, (c) determine resource allocation, and (d) validate organizational credibility. Other roles of program evaluation found from the present study also corresponded with "process" (Ebrahim, 2010, p. 107) accountability efforts: (a) advance organizational learning, (b) inform program management decisions, (c) cultivate funding collaborations, and (d) enhance communication with multiple stakeholders.

Ebrahim's (2010) "hows" of organizational accountability were similar to Thomas's (2010) prescription for nonprofits to provide information on outcomes as a form of accountability. Thomas noted, "Nonprofit organizations need, at a minimum, to engage in systematic outcome assessment—that is, regular measurement and monitoring of how well their programs are performing relative to the desired outcomes" (pp.401-402). The results from the present study regarding the role of program evaluation found that assessing impact and determining outcomes was the highest ranked purpose for program evaluation.

Inclusive to the concept of organizational accountability is organizational effectiveness. Organizational effectiveness is most often determined by the degree to which an organization achieves identified goals (Murray, 2010). Ultimately, the mission of a nonprofit is the chief goal for the organization (Worth, 2009). Accordingly, the findings from the present study reinforced the role of program evaluation as a means to ascertain impact, effectiveness, and mission alignment.

Information gained from the present study regarding the role of program evaluation can inform an overall organizational accountability framework. The multiple roles of program evaluation found in the present study may provide part of the "hows" for nonprofit accountability frameworks such as posited by Candler and Dumont (2010).

An interesting finding of the present study was the amount of emphasis placed on organizational learning through reflective practice. Organizational learning ranked second highest on the Delphi surveys, and reflective practice was a significant topic of conversation for those leaders interviewed. However, previous studies conducted with professional evaluators (Lee et al., 2008) and nonprofit leaders (Carman & Frederick, 2008) did not note significant emphasis on program evaluation as a tool for advancing organizational learning through reflective practice. Yet Patton (1999, 2011) posited that, in particular, action research and developmental evaluation approaches were well suited for reflective practice and organizational learning.

The multi-year reflective practice project sponsored by the Jessie Ball duPont Foundation and the Community Foundation underscored the value of using program evaluation as a tool for advancing organizational learning through reflective practice. The significance of this five-year project for the Northeast Florida nonprofit sector cannot be understated. From these efforts, reflective practice became a framework for formative evaluation efforts in the sector. Additionally, reflective practice became the standard for processing information from other evaluative efforts. Moreover, findings suggested that the culture of the Northeast Florida nonprofit sector became more collaborative and open to program evaluation efforts due to the reflective practice initiative. The weaknesses of the initiative were that it was not implemented sector-wide, it was expensive, and it only lasted five years. However, in the recommendations section of this chapter, I will discuss how the sector can build on all of the advantages gained from the reflective practice initiative.

Funder and Provider Collaboration

Participants' perceptions regarding the relationships between providers and funders varied. Four main areas of discussion emerged: (a) distribution of power, (b) funder requirements, (c) communication and collaborative efforts, and (d) program evaluation culture. These results are similar to a review of the literature that indicated how the different types of relationships between the two groups influence program evaluation efforts.

First, findings from the present study indicated that the relationship between providers and funders represented an unequal distribution of power particularly regarding access to needed resources (e.g., money, capital). Interview data from both providers and funders supported the fact that funders control access to financial resources. These results affirm resource dependency theory. As noted previously, resource dependency theory recognizes that one group (providers) is beholden to another group (funders) for their existence (Carman, 2011). Providers are dependent on funders for resources so they use program evaluation strategies to enhance their efforts to acquire funding and other resources. Results from the present study found that part of the role of program evaluation is to cultivate funding collaborations and enhance communication with multiple stakeholders as tactics to garner needed resources. Rather than being equal partners, results indicated an unequal status between the two groups that fosters dependency.

The second theory relevant to the findings of the present study is agency theory. As noted previously, agency theory describes the roles between two parties as either principals or agents. In short, agents do the bidding of the principals and trust between the two is sometimes lacking. The role of the principals is to ensure that agents are performing as expected (Carman, 2010, 2011; J. L. Miller, 2002). The results of the present study seem to suggest that some participants viewed the relationship between providers (agents) and funders (principals) as an example of agency theory. This is evident through the detailed information from the present study regarding the multiple requirements funders have regarding program evaluation efforts.

The results of the present study indicated an emphasis on the role of program evaluation as providing information related to outputs, effectiveness, outcomes, and quality assurance measures. Funders are trying to ensure that providers are doing what they are paid to do. For example, a Delphi panelist concluded, “I consider this [program evaluation] very important in showing accountability to funders and the community.”

The third theory relevant to the results of the present study regarding the relationship between providers and funders is stewardship theory. Stewardship theory suggests that the parties involved are partners dedicated to a common purpose. Efforts focus on improving services and moving the shared agenda forward (Carman, 2010, 2011). The results for the present study indicated that funders, in particular, and some providers regarded their relationship in this manner. This is demonstrated through the emphasis on reflective practice efforts as reported by study participants, particularly funders. It is important to note that all funders interviewed focused their attention on how they could develop stronger and more productive relationships with providers focused on improving services for constituents. Of course, the funders in the study realized that they had more access to resources. Yet their language centered on how to build partnerships, how to make change together, and an understanding that funders are not successful if the provider organizations are not successful.

The Capacity for Program Evaluation

The need for and lack of funding for program evaluation efforts were common findings of the present study. The need for funding ranked as the second highest element of capacity required but ranked lowest as to the current level of funding available for program evaluation efforts. This is not surprising. A review of the literature indicated that funders were often reluctant to provide the needed resources (Carman, 2010). As an example, I am involved with a funding collaboration in Northeast Florida. As part of the organization, I reviewed grant proposals for a competitive grant process. One of the requirements for the grant was that providers must have an evaluation component. However, the funding organization will not fund those efforts. I provided feedback regarding the discrepancy between the requirement and the lack of funding. I also suggested that funds be available above the amount requested for evaluative purposes. Unfortunately, this did not occur. In addition to funders putting resources into evaluation, nonprofits also rarely invest their own resources in program evaluation efforts (Carman, 2007). In fact, study participants could only identify two organizations in the Northeast Florida nonprofit sector that had staff dedicated to program evaluation efforts.

Another essential capacity element result also discussed in the literature is the importance of a positive culture regarding program evaluation efforts. Results from the Delphi surveys and interviews found that having a positive culture was necessary for quality program evaluation strategies. Results of the present study parallel previous information found in a review of the literature that a positive program evaluation culture advanced program evaluation efforts. The literature reviewed for the present study reported that many elements of organizational culture and sector-wide culture have an effect on program evaluation efforts (Behrens & Kelly, 2008; Bozzo, 2002; Stevenson et al., 2002).

Many nonprofit organizations lack the time for evaluation efforts (Carman & Fredericks, 2009; Hendricks et al., 2008; Kegeles et al., 2005). The results of the Delphi surveys phase of study also supported the finding that time was an essential element needed for program evaluation capacity. Panelists rated time as the highest ranked capacity element needed. Yet panelists also recognized that sufficient time was lacking to implement quality program evaluation strategies.

Finally, the relationship between providers and funders was captured in the categories that focused on ongoing collaboration in the nonprofit sector and the need for funders to have realistic expectations regarding program evaluation efforts. In fact, as interviewees noted, many of the aforementioned capacity elements can be addressed if collaboration increased. Improving collaboration around program evaluation efforts between providers and funders has been previoulsy reported to be an important element to capacity building (Atkinson et al., 2005; Bozzo, 2002; Carman, 2010, 2011).

Ultimately, the results from the present study were analogous to other research regarding the needed capacity elements for program evaluation. Findings from the present study reflected in a review of the literature included the need for time, money, skilled people, collaboration, and a culture ripe for program evaluation (Carman & Fredericks, 2009; Duignan, 2003; Hendricks et al., 2008; Kegeles et al., 2005; Kehrer, 1993). Yet the final order of rankings from the category with the overall highest mean (sufficient time) to the category with the overall lowest mean (ongoing training) of the essential program evaluation capacity elements were particular to the perceptions of the participants from the present study.

The results from the present study included some considerations regarding the development of program evaluation capacity. The need to find a sector-wide champion to promote program evaluation and capacity building efforts was a finding of the present study. Bozzo (2002) also suggested having a group of leaders in a nonprofit sector from various roles (e.g., funders, program evaluators, service providers) unify to advance the cause of program evaluation. A review of the literature also noted the need for a program evaluation champion or advocate within individual organizations (Garcia-Iriarte et al., 2011; Kegeles et al., 2005; Milstein et al., 2002). The strategy of the sector-wide champion from the present study seemed particularly important to the Northeast Florida nonprofit sector as it is similar to how the reflective practice initiative was implemented and maintained for so many years. Two funding agencies took the lead on this project as well as investing a great deal of time and money into the effort. The results indicated that this type of champion, either a person or organization, is the first step to building program evaluation capacity in Northeast Florida.

The finding from the present study regarding the importance of collaboration with sector stakeholders, institutions of higher education, and others as a program evaluation capacity-building component were similar to suggestions found in a review of the literature (Arnold, 2006; Atkinson et al., 2005; Compton et al., 2002; T. I. Miller et al., 2006; Stevenson et al., 2002;).

In summary, the findings from the present study regarding the role of and capacity for program evaluation reinforced the conceptual framework for the study (see Figure 1). The overall conceptual framework included program evaluation as an element of organizational accountability. Furthermore, organizational accountability supports the mission of nonprofits. The other part of the conceptual framework for the present study was how the collaboration between providers and funders influences both program evaluation efforts and evaluation capacity. Additionally, evaluation capacity has direct bearing on program evaluation efforts. The results of the present study indicated the interdependent influences among these three aspects (collaboration, evaluation capacity, and program evaluation) of the conceptual framework.

Limitations of the Study

The intent of the present study was to understand the perceptions of service providers and funders regarding the role of and capacity for program evaluation. Particularly, the nonprofit sector in one area, Northeast Florida, was the focus of the present study. Furthermore, the participants chosen for the study were an elite group of leaders in the sector. As mentioned previously, generalizability was not a goal of the study. However, as Donmoyer (1990) suggested, others may find heuristic value in the results from the present study, particularly those from the nonprofit sector. However, aspects of the sampling strategies and survey construction presented limitations.

The sampling strategy used to garner participation created limitations. As the researcher, I created a list of possible provider participants from my connoisseurship of leaders in the Northeast Florida sector who met the criteria. Consequently, the list was limited by my knowledge and contacts. Additionally, I used referrals and snowball sampling techniques to secure funder participants. As a result, the funder participant pool was limited to private funding organizations in the sector. These approaches for acquiring program participants also limited the ethnicity and gender distribution of participants. Furthermore, I did not include anyone representing government funding, which limited the findings related to funders' perceptions to those in the private sector.

The Delphi survey instruments also had limitations. The construction of the second and third Delphi survey instruments could have been clearer and designed better. At the time I constructed the instruments for these rounds, I could not figure out how to incorporate a section for comments with each rated category. The design for the final Delphi round instrument was much better than the previous rounds, as I discovered how to incorporate a comments section with each rated category. As a result, I received much more feedback from panelists.

The phrasing of the instructions for rating the essential capacity elements section may have been a limitation of the present study. As mentioned previously, at least one Delphi panelist misinterpreted the intent of that particular section for the second Delphi survey round. Although the findings did not indicate that others misinterpreted the instructions, one cannot be sure. Clearer phrasing might have produced different results.

As mentioned in the previous chapter, only choosing six of the categories after the second Delphi survey round to use for ratings for the third Delphi survey round was shortsighted. Because the final ratings were not available prior to the third Delphi survey round, I should have included all essential capacity elements. Because I omitted two of the categories, I have limited data about the perceptions of the panelists regarding the current level of capacity for on-going training and realistic expectations of the philanthropic community.

All research has limitations. However, what is interesting to note is that the limitations of the present study parallel some of the capacity challenges associated with program evaluation efforts such as skilled human resources, methodological issues, expectations, and ongoing training. Furthermore, as the results for the role of program evaluation suggested, these limitations can become the springboard for learning. Consequently, I learned valuable lessons regarding research and would make different choices for future research.

Major Conclusions

The results of the present study produced information regarding the role of and capacity for program evaluation. Several major conclusions came from an analysis of the study as a whole. The core conclusions concentrated on the interest of the subject, reflective practice, program evaluation efforts, funder-driven priorities, and capacity development.

Interest

The topic of program evaluation and the related capacity issues appeared to be of great interest to participants. The interest is important for several reasons. First, interest in the topic and commitment to the sector kept panelists engaged throughout the entire four rounds of the Delphi survey process. Additionally, every person interviewed reported interest in the subject as well as agreement with the categories produced from the Delphi surveys. Although the focus of the present study was the perceptions of leaders regarding the role of and capacity for program evaluation, their collective commitment to the project and expressed interest cannot be understated.

The topic of program evaluation is not a new one for these leaders. Discussion regarding impact, outcomes, outputs, and other metrics has been consistent through the years. However, most of the discussions have been between specific funders with specific providers. The difference in interest with the present study is the sector-wide examination of the broader aspects of the role of and capacity for program evaluation. The study provided leaders the opportunity to consider mechanisms for building evaluation capacity. Moreover, the study also provided the venue to highlight positive sector initiatives such as reflective practice and collaboration efforts. As one Delphi panelist concluded, “Glad to participate [in the study]. I hope some useful information comes out of the present study to assist nonprofits with incorporating program evaluation into our agencies on a more consistent basis.”

Reflective Practice and Organizational Learning

Results of the present study indicated the preeminence of reflective practice as a vital factor of program evaluation efforts in the Northeast Florida nonprofit sector. Study participants viewed reflective practice as part of a role of program evaluation as well as a method of evaluation. This is consistent with Patton’s (1999, 2011) use of reflective practice as a part of evaluation designs where information was needed in real time to make decisions regarding program development.

In particular, the funders interviewed and those on the Delphi panel were very vocal about the benefits of reflective practice. They viewed reflective practice as an avenue to advance organizational learning, improve services, and to increase collaborative efforts in the sector. The importance of reflective practice for organizational and systems-wide change is consistent with the literature. Ebrahim (2010) described the process of using reflective practice as "adaptive learning in which nonprofits create regular opportunities for critical reflection and analysis in order to make progress toward achieving their missions" (p. 113) and listed this as a key means of internal organizational accountability.

Although study participants viewed reflective practice as a key aspect of organizational learning, their discussion of organizational learning appeared limited and did not include all of the required elements (personal mastery, mental models, shared vision, team learning, and systems thinking) suggested by Senge (1990) that support organizational learning. Of these five elements, the use of reflective practice in the Northeast Florida nonprofit sector appeared to promote two of Senge's elements: team learning and mental models. Study participants reported using reflective practice to challenge their assumptions regarding what they knew about a particular issue or practice and used this information to advance change. In addition, reflective practice was used to develop team learning within organizations regarding service delivery. Funders and providers used reflective practice to discover ways to improve services in order to achieve the desired outcomes of the funded program. This mirrors Senge's definition of team learning, which is "the process of aligning and developing the capacity of a team to create the results its members truly desire" (Senge, 1990, p. 218). To a limited degree, study participants reported using reflective practice as a means promote systems thinking in regards to how funders and providers collaborate with each other. However, using reflective practice for personal growth and creating a shared vision for the sector was not explicitly detailed by study participants.

Reflective practice is also a key element for evaluation capacity building. Several of the leaders interviewed commented on how creating a learning environment between providers and funders increased collaboration. Furthermore, the perceptions of study participants indicated that increased collaboration between providers and funders might lead to increasing other capacity elements such as time, money, and realistic expectations from funders. The literature also suggested that reflective practice is foundational to the work of evaluation capacity building (Baizerman et al., 2002).

Focus of Program Evaluation Efforts

The results of the present study provided eight distinct yet overlapping categories regarding the role of program evaluation. Delphi survey results found nuances among the categories, which were more difficult to determine from interviews alone. The highest ranked category with an overall mean of 4.83 was a focus on the role of program evaluation as a way to measure impact, program effectiveness, and outcomes. The lowest ranked category of the role of program evaluation had an overall mean of 3.92. This category was related to validating organizational credibility through mission alignment and strategic planning.

The lowest rating for this category is of particular interest on two levels. First, the only reason a nonprofit exists is to carry out the mission of the organization, which should also be a public benefit (Worth, 2009). Furthermore, a review of the literature indicated that the executive leader of a nonprofit has the primary duty to focus on the mission of the organization (Herman, 2010; Worth, 2009). It would seem, therefore, that executive leaders would want information obtained from program evaluation efforts that confirm mission alignment, or conversely, indicate mission drift. This information would be highly valuable for future planning, including course corrections, if needed. Yet the perceptions of the Delphi panelists from the present study placed mission validation/alignment as the lowest ranked role of program evaluation.

Second, the low emphasis on the use of program evaluation for strategic planning purposes is also telling. Unfortunately, too often the idea of strategic planning is only a SWOT (strengths, weaknesses, opportunities, threats) analysis to create a list of activities to accomplish. At best, this type of strategic planning becomes a checklist of accomplishments. At worst, the plan sits on somebody's shelf or in somebody's computer to be rolled out when requested by funders. Either of these scenarios limits the vital role of strategic planning in the overall health of an organization.

Bryson (2004) defined strategic planning as "a disciplined effort to produce fundamental decisions and actions that shape and guide what an organization (or other entity) is, what it does, and why it does it" (p. 6). Strategic planning is also a means to refine, promote, and support the mission of an organization. Strategic planning also includes a feedback loop for assessing results (outcomes, goals, and objectives) and strategies (Bryson 2004, 2010). Correspondingly, program evaluation efforts should be a prime part of the strategic planning process. However, the use of program evaluation effort as a tool for strategic planning had the lowest overall rating from the Delphi survey results. Consequently, the low priority placed on program evaluation for strategic planning from the results of the Delphi surveys reinforced the perception that leaders have not instituted program evaluation efforts as an integral part of organizations.

Funder Driven Priorities

The results of the present study indicated that the funders in the Northeast Florida nonprofit sector drive the efforts regarding program evaluation. Funders set the priority focus areas for evaluation (e.g., impact, outcomes). Funders provide funding for efforts they value (e.g., reflective practice). Finally, funders decide to what degree evaluation efforts are supported through distribution of funds, technical assistance, and other capacity-building strategies.

The power of funders to create a program evaluation culture in the Northeast Florida nonprofit culture was evident from the results of the present study. Both providers and funders reported that most of the program evaluation efforts in the sector stem from funder requirements. Consequently, the nonprofit culture was initially resistant to program evaluation. The negativity was further exacerbated by the different funder demands and lack of willingness from the funding community to pay for the evaluation requirements. Essentially, program evaluation was forced on the nonprofit community as an unfunded mandate.

Funders also had the power to change the culture of the nonprofit community and the relationship between providers and funders. They took on the challenge by funding and participating in the reflective practice initiative. The results of the five-year project produced greater collaboration within the Northeast Florida nonprofit sector, a greater appreciation for reflective practice, and, according to study participants, was the genesis of the Nonprofit Center of Northeast Florida. The Nonprofit Center of Northeast Florida's role is to continue the work of the reflective practice project through "connecting, strengthening, and advocating for nonprofits to create a more vibrant Northeast Florida" (Nonprofit Center of Northeast Florida, 2013a).

Capacity Development

The results of the present study produced several capacity elements needed for quality program evaluation. Results of the present study also indicated that study participants perceived that the Northeast Florida nonprofit sector has a deficiency of program evaluation capacity. Unfortunately, the lack of program evaluation capacity was typical of findings from a review of the literature (Carman, 2007, 2009).

Findings from the present study produced very clear steps that study participants perceived could increase program evaluation capacity in Northeast Florida. The steps included identifying a leader for the effort, educating the sector on program evaluation and the associated costs, studying model communities, and building collaborations with local universities and colleges. This plan for increasing capacity seems feasible, and the time may be right for implementation as demonstrated by the interest in the topic.

Recommendations

The role of and capacity for program evaluation in the nonprofit sector is an important topic of study. The ability for any individual organization or the sector as a whole to institute program evaluation strategies and build capacity is directly influenced by the leadership of the sector. Accordingly, the present study focused on the perceptions of leaders in the Northeast Florida nonprofit sector regarding these multi-faceted variables.

As an exploratory study, results regarding the role of and capacity for program evaluation generated several recommendations for future practice and research. In fact, these recommendations provide input regarding the reflective practice question: What are we going to do with this information? To answer this question, I have several recommendations on how to advance the practice of program evaluation in the nonprofit sector and how to continue research on the subject.

Recommendations for Practice

Although the present study was not designed for generalization, the findings may be of interest to other providers and funders in other nonprofit sectors. The results of the present study support four major recommendations for practice. First, I recommend that leaders of nonprofit organizations examine how the different roles of program evaluation can serve their organizations. Rather than just concentrating program evaluation efforts on producing reports, documenting impact, or garnering additional revenue, that nonprofit leaders should incorporate program evaluation efforts into every aspect of their work, particularly strategic planning.

Second, I recommend that nonprofit leaders invest their resources into program evaluation efforts. As an integral part of strategic planning and mission advancement, program evaluation efforts affect every aspect of an organization. Consequently, allotment of organizational resources should be comparable to those provided to financial and development departments.

Third, the Northeast Florida nonprofit sector as a whole should prioritize the development of increased program evaluation capacity. The results from the present study provide ideas for increasing capacity. It is imperative that an organization take on the role as a leader to further these efforts. Results from the present study suggested that the Nonprofit Center of Northeast Florida would be a logical choice as a sector-wide champion of this effort. I recommend that the Nonprofit Center of Northeast Florida create a workgroup of providers and funders to explore implementation of the steps suggested from the present study. One of the first steps for the lead person or organization of this effort is to engage the research and evaluation communities from the local colleges and universities. As an important part of this recommendation, the workgroup should also seek models of provider/funder relationships that exemplify program evaluation capacity building, integration of program evaluation in strategic planning efforts, and organizational leadership support of program evaluation.

Fourth, I recommend that the Northeast Florida nonprofit sector build on the reflective practice initiative that started over a decade ago by involving more nonprofits and by strategically integrating the multiple variables of organizational learning suggested by Senge (1990) into the scope of the initiative . Since the ending of the project, leadership in the sector has changed, new nonprofits have been created, and nonprofit leaders have focused their attention on surviving after the 2008 financial crisis. Consequently, some of the gains created from the reflective practice initiative may have diminished. This was important work to the sector, which created strong collaborations. If history is a predictor, the results can provide substantial benefit to the Northeast Florida nonprofit sector.

Finally, due to the expressed interest from leaders involved in the study, I recommend that the results of the present study be redacted into an executive summary for distribution to study participants and other leaders in the Northeast Florida nonprofit sector.

Recommendations for Research

The purpose of the present study was to explore the perceptions of nonprofit and funding leaders regarding the role of and capacity for program evaluation in the Northeast Florida nonprofit sector. The present study was limited to an elite group of leaders within the Northeast Florida nonprofit sector. Accordingly, the information obtained was not meant to be predictive, inferential, or even representative of other nonprofit sectors. Exploratory studies of the nature of the present study often produce findings that lead to other types of research with quantitative, qualitative, and mixed designs. Correspondingly, the results from the present study provided information that could be the springboard for future research.

First, I would recommend conducting research similar to the present study with a broader base of participants. Further research could include a greater variety of nonprofit organizations, public (government) funders, board members, individual donors, and frontline staff doing the work. Additionally, I would be sure to include participants with a greater degree of diversity as it relates to race and gender. It would be interesting to analyze the results with a broader base of representation.

Second, I recommend additional research regarding the eight categories for the role of program evaluation that came from the Delphi surveys. I would suggest using the categories as a basis for a quantitative survey-based study. I would recommend sending the survey to a larger random sample of nonprofit leaders (providers and funders) in Northeast Florida or a wider geographical area. It would be interesting to use quantitative analysis to examine the differences between sub-group representation (provider or funders), years of experience, and other relative demographic information. I would recommend using the eight categories representing the essential capacity elements for a study in a similar fashion.

Third, I recommend that a new study focus on program evaluation capacity building. It would be interesting to design a quantitative pre/post intervention study with pre-determined metrics to rate the level of program evaluation capacity in a nonprofit sector. The pre-intervention instrument could ascertain current levels of program evaluation capacity. After obtaining the results from the pre-intervention instrument, I would recommend implementing the capacity building steps identified in the present study. After this intervention (implementing the designated capacity building steps for a pre-determined amount of time), I would use the instrument for post-intervention data to measure the levels of program evaluation capacity. This type of research would take time and money but could provide valuable knowledge for increasing program evaluation capacity.

Fourth, I recommend further research that examines why funders are reluctant to invest in program evaluation or program evaluation capacity building. A study of this nature could also include exploring the inconsistencies between the perceived resources needed for program evaluation and the reluctance to provide funding for these resources.

Fifth, although not a main point of the present study, I recommend further research regarding the specific uses of different program evaluation methods in the nonprofit sector. A study of this nature could help determine the level of knowledge that people in the nonprofit sector have regarding the various program evaluation approaches, implementation strategies, and analysis methods.

Sixth, further research could explore how widespread the use of reflective practice is among all of the nonprofits in the Northeast Florida nonprofit sector. The entire nonprofit sector was not involved in the five-year reflective practice initiative. Consequently, there cannot be an assumption that the knowledge and advances made from the initiative filtered to other nonprofit organizations in the sector. Additional research could provide some clarity regarding the depth and breadth the impact of the reflective practice initiative had in the sector as a whole.

Seventh, I recommend future research regarding the extent to which reflective practice, organizational learning, systems thinking, and program evaluation efforts impact organizational effectiveness. A part of the research could include comparing nonprofit organizations that participate in organizational learning and capacity-building initiatives with those who do not.

Finally, I recommend additional research solely focused on defining the relationship between providers and funders such as the attributes and dimensions of funder/provider partnerships. The results of the present study indicated the importance of the dynamics between the two groups regarding program evaluation and capacity building efforts. However, another study could focus this on a broader scale by examining how funder/provider collaborations influence the overall culture, productivity, and cooperation of any given nonprofit sector.

Conclusion

Although there has been research on program evaluation efforts in the nonprofit sector, the present study is unique in that it explores the subject matter from the perspectives of human services provider nonprofits and funders from the Northeast Florida nonprofit sector. Also unique to the present study was the use of the Delphi surveys as a means for a nominal group process. The Delphi surveys phase of the study allowed for a group process while protecting the confidentiality of the participants. This was important as providers needed to provide input without fear of retaliation from funding sources and funders needed the opportunity to speak freely without fear of peer pressure or other negative feedback. Additionally, the qualitative design of the study provided the opportunity for participants to generate their own categories relevant to the study. Furthermore, interviews provided an opportunity for the expression of voice of the unique perspectives for both providers and funders.

The implementation of quality program evaluation is hindered when capacity is lacking. Yet the interest generated from the present study indicated that this is a prime time to advance the cause of program evaluation and capacity building efforts. The results of the present study can serve as a catalyst for increasing capacity. Leaders involved with the present study mapped out a plan for creating capacity. The question remains as to whether these leaders will follow through with the time and money required for these efforts to possibly increase program evaluation capacity.

The Northeast Florida nonprofit sector has a clear advantage with the level of cooperation so evident in the interactions among providers, funders, and the larger community. Study participants emphasized how this culture of collaboration is unique to Northeast Florida. Leaders of the sector could leverage the collaborative nature of the sector to engage government officials, businesses, and the community to support the development of program evaluation capacity as a whole.

Conclusions are often thought of as endings. Yet rather than being the end, the information from the present study can serve as a foundation for future research, practices, and community building. I think that it is only fitting that the final word regarding the role of and capacity for program evaluation come from a study participant. A senior executive with a funding organization observed, “This is intentional work, you gotta plan for it, you gotta fund it, you gotta be able to use what you learn. It can’t sit on a shelf.”

Appendix A

IRS Tax Exempt Categories

- Corporations organized under an Act of Congress for use by the United States.
- Corporations created to hold title to property where any income collected is given to another organization exempted under section 501(c).
- Organizations established for religious, charitable, scientific, public safety testing, literary, educational, fostering international amateur sports competitions, and prevention of abuse to animals or children purposes. This category of exempt organizations also includes private foundations. Organizations exempted in this category are commonly known as 501(c)3's.
- Civic leagues, social welfare institutions, or employee associations.
- Labor, agriculture, or horticulture organizations.
- Business leagues, chambers of commerce, real-estate boards, boards of trade, or not for profit professional football leagues.
- Domestic fraternal societies, orders, or associations.
- Local teachers' retirement fund associations.
- Local benevolent life insurance associations and cooperative utilities companies.
- Not for profit cemetery companies.
- Not for profit credit unions, corporations, or associations without capital stock or organized for the purpose of providing reserve funds for associations or cooperative/mutual banks.
- Insurance companies with limited revenue and mutual insurance companies.
- Cooperative organizations created for the purpose of financing crops.
- Trusts created to provide supplemental unemployment benefits.
- Trusts created to provide pension benefits.

- Organizations, posts, trusts, foundations, auxiliary units with current and previous members from the United States Armed Forces without individual shareholder profits.
- Nonprofit legal aid organizations.
- Black lung trusts.
- Trusts created to pay related costs for employee retirement plans.
- United States Armed Forces associations created before 1880.
- An organization created to hold title to property where any income is given to multiple organizations exempted under section 501(c).
- Not for profit organizations created by a state that provide health insurance to the uninsured.
- Organizations created by a state that provide workmen’s compensation insurance.
- The National Railroad Retirement Investment Trust established under section 15(j) of the Railroad Retirement Act of 1974.
- Religious and apostolic organizations.
- Cooperative hospital service organizations.
- Cooperative service organizations of operating educational organizations.

Appendix B

Interview Questions

1. Please discuss your reflections (e.g., surprises, clarifications, concerns, expansion of ideas) from the findings of the Delphi survey.

    a. Role of:
    b. Essential Capacity Elements:
    c. Reality Check:

2. What is the role of (e.g., meaning, purposes, benefits, approaches, utilization) program evaluation in the Northeast Florida nonprofit sector?

3. How do the relationships between service providers and funders influence program evaluation efforts in the Northeast Florida Sector?

4. What are ways that the Northeast nonprofit sector can develop program evaluation capacity?

5. What else would you like to say that you have not had a chance to share?

Appendix C

Interview Informed Consent

Dear Participant:

My name is Gail Patin and I am a student in the Educational Leadership (Ed.D.) doctoral program at the University of North Florida (UNF). I am conducting a research study on program evaluation efforts in the nonprofit sector. This study aims to learn the perceptions of leaders in funding organizations and human services agencies regarding the role of and capacity for program evaluation in the Northeast Florida nonprofit sector.

I invite you to take part in this study as you are identified as a leader expert in a nonprofit human services or funding organization. You were selected for an interviews because of need for clarification from survey results, interest in the subject, an expressed interest in being interviewed, or you were reccomended for an interview by others in the nonprofit sector. You will be asked to take part in an interview that will last approximately 45 minutes to an hour. No one other than myself will know your identity and I will maintain your individual responses with the strictest confidentiality. I will not share your name, the name of your organization or other identifying information. I will record the interviews using a digital recorder(s). I will download the audio digital files and store on a UNF secure server (e.g., UNF Osprey h drive, UNF Osprey Skydrive). Once transferred, I will delete the recordings from the recorder. I will maintain the digital recordings of the interview on a UNF secure server until I have completed the study. After that point, I will destroy all digital recordings through permanent deletion. Other electronic copies of data (transcripts, notes, etc.) will also be stored on a UNF secure server (e.g., UNF Osprey h drive, UNF Osprey Skydrive).

I will transcribe and/or use transcriber(s) in order to have written transcripts. I will require transcribers to sign a confidentiality agreement prohibiting them from disclosing any of the information to others. Since I may use transcribers, I will provide the transcriber(s) an electronic copy of the interview. I will deliver the recording in person through a portable hard drive for the transcriber to copy onto his/her computer. After the transcriber completes the document, he/she will destroy through permanently deleting his/her copy of the recording. Data gathered from the interview will be confidential to the extent allowable by law.

As a direct benefit for taking place in the interview, I will provide you with a final copy of the Delphi results from the early part of the study. Additionally, others may benefit from the information we learn from the results of this study. However, you will not be compensated for your participation. There are no foreseeable risks for taking part in this study. Participation is voluntary and there are no penalties for skipping questions or withdrawing your participation. Thus, you may choose to withdraw from this study at any time with no penalty or loss of benefits you would otherwise be entitled to receive.

If you have any questions or concerns about this study, please contact me or my professor, Dr. Katherine Kasten. If you have questions about your rights as a participant, you may contact the University of North Florida's Institutional Review Board Vice Chairperson, Dr. Krista Paulson, I thank you for your consideration.

Sincerely,

I______________________________(print name) attest that I am at least 18 years of age and agree to take part in this study. A copy of this form was given to me to keep for my records.

Signature:_______________________________________Date:__________________________

Appendix D

First Delphi Round Survey

Default Question Block

Dear Participant:

My name is Gail Patin and I am a student in the Educational Leadership (Ed.D.) doctoral program at the University of North Florida (UNF). I am conducting a research study on program evaluation efforts in the nonprofit sector. This study aims to learn the perceptions of leaders in funding organizations and human services agencies regarding the role of and capacity for program evaluation in the Northeast Florida nonprofit sector.

I invite you to take part in this study as you are identified as a leader expert in a nonprofit or funding organization. If you take part in my study, you will be asked to respond to an online multi-round Delphi survey regarding the role of and capacity for program evaluation in the nonprofit sector. The Delphi Survey consists of two to four (2-4) rounds. The first questionnaire is exploratory in nature where I, as the researcher, ask for your expert opinion on a series of broad questions. The first survey is estimated to take between 20-40 minutes. Subsequent surveys are expected to take half the amount of time as you will be rating and/or priority ranking information I condensed from the initial questionnaire. Data gathered from the Delphi survey will be confidential to the extent allowable by law. No one other than myself will know your identity and I will maintain your individual responses with the strictest confidentiality. I will not share your name, the name of your organization or other identifying information.

As a direct benefit for taking place in the Delphi survey, I will provide you with a final copy of the Delphi results. Additionally, others may benefit from the information we learn from the results of this study. However, you will not be compensated for your participation. There are no foreseeable risks for taking part in this study. Participation is voluntary and there are no penalties for skipping questions or withdrawing your participation. Thus, you may choose to withdraw from this study at any time with no penalty or loss of benefits you would otherwise be entitled to receive. Additionally, I may contact you for an interview following the Delphi survey of the interview. Participation in the interview is also voluntary.

If you have any questions or concerns about this study, please contact me or my professor, Dr. Katherine Kasten. If you have questions about your rights as a participant, you may contact the University of North Florida's Institutional Review Board Vice Chairperson, Dr. Krista Paulson, at I thank you for your consideration.

Sincerely,

To print this informed consent page, look for a print icon at the top of the web browser or under the "File" drop down menu.

- ○ By clicking here, I confirm that I am at least 18 years of age and understand and accept the above information. I agree and consent to participate in this study.
- ○ No, I do not wish to participate in this study.

Thank you for agreeing to participate in this study on the role of evaluation in the nonprofit sector of Northeast Florida. This survey consists of 8 questions.

Here are some tips that may help you navigate this survey.

1. This survey may be completed in more than one session. In order to do so, please follow the steps listed below so the responses you put in will be saved.

- **Be sure to save the email I sent as it contains the link that will take you back to your survey.**
- **After you enter your responses for any given question, you must click the "Save/Next" button or the "Previous" button in order to save your responses. If you close out of the program before doing this, your responses will not be saved.**
- **After saving your responses, as noted above, you may simply close the web tab/page by clicking on the "X" button (usually in the top right hand corner of a web tab/or page) to leave the survey.**
- **Use the link in the email to return to your saved survey responses.**

2. Do not click the "Finish" button until you are completely finished and ready to submit your responses.

3. If you have any questions/issues/concerns, do not hesitate to contact me (Gail Patin)
You may contact me at anytime (even early mornings or later in the evening). I want to make this experience as easy as possible for you.

I appreciate your willingness to share your thoughts, expertise and insight.

1. When you hear or read the phrase, *program evaluation*, what does this mean to you? In a short phrase or sentence, please describe *at least three* (more are appreciated) thoughts that come to mind.

1

2

3

4

5

2. What are the purposes for conducting program evaluations in nonprofit organizations? In a short phrase or sentence, please describe *at least three* (more are appreciated) thoughts that come to mind.

1

2

3

4

5

3. What are the benefits to nonprofit organizations for conducting program evaluations? In a short phrase or sentence, please describe *at least three* (more are appreciated) thoughts that come to mind.

1

2

3

4

5

4. What are some program evaluation approaches used in the Northeast Florida nonprofit sector? In a short phrase or sentence, please describe *at least three* (more are appreciated) thoughts that come to mind.

1

2

3

4

5

5. How can nonprofit organizations use program evaluation information collected? In a short phrase or sentence, please describe *at least three* (more are appreciated) thoughts that come to mind.

1

2

3

4

5

6. What are main challenges the nonprofit sector in Northeast Florida faces regarding program evaluation? In a short phrase or sentence, please describe *at least three* (more are appreciated) thoughts that come to mind.

1

2

3

4

5

7. What (if any) are the resources needed to develop capacity in order to sustain program evaluation efforts in the Northeast Florida nonprofit sector? In a short phrase or sentence, please describe *at least three* (more are appreciated) thoughts that come to mind.

Click to write Choice 1

Click to write Choice 2

Click to write Choice 3

Click to write Choice 4

Click to write Choice 5

8. How do funders and providers in the Northeast Florida nonprofit sector work together to develop program evaluation strategies (objectives, outcomes, evaluation purposes, uses of evaluation results, evaluation approaches, capacity issues, resource allocations, etc…)? In a short phrase or sentence, please describe *at least three* (more are appreciated) thoughts that come to mind.

Click to write Choice 1

Click to write Choice 2

Click to write Choice 3

Click to write Choice 4

Click to write Choice 5

PLEASE STOP & READ

If you are <u>NOT FINISHED</u> with the survey and wish to come back later to complete, <u>DO NOT</u> click on the "Finish" button below.

Instead, simply close the web tab/page by clicking on the "X" button (usually in the top right hand corner of a web tab/or page) to leave the survey. Use the link in the email to return to your saved survey responses. This will save your responses for when you return to complete the survey.

If you have any questions/issues/concerns, do not hesitate to contact me (Gail Patin) at

You may contact me at anytime (even early mornings or later in the evening). I want to make this experience as easy as possible for you.

If you have completed the survey, then click on the "Finish" button below.

Appendix E

Approval Emails from Authors

Gail Patin

FW: 2008 Evaluation and Program Planning article: Practices and challenges in educational program evaluation in the Asia-Pacific region: Results of a Delphi Study

1 message

Gail Sun, Apr 14, 2013 at 5:39 PM
To:

——

From: james altschuld
Sent: Tuesday, July 17, 2012 11:55 PM
To: 'Gail'
Cc: 'Yi-Fang Lee'; 'Hung, Sonya'
Subject: RE: 2008 Evaluation and Program Planning article: Practices and challenges in educational program evaluation in the Asia-Pacific region: Results of a Delphi Study

Best always. JWA

From: Gail
Sent: Tuesday, July 17, 2012 11:30 PM
To: 'james altschuld'
Cc: 'Yi-Fang Lee'; 'Hung, Sonya'
Subject: RE: 2008 Evaluation and Program Planning article: Practices and challenges in educational program evaluation in the Asia-Pacific region: Results of a Delphi Study

Thank you everyone for allowing me to use your questions. I certainly will cite you in my references. Also thank you for responding so quickly. I really do appreciate it.

Peace,

Gail Patin

——

From: james altschuld
Sent: Monday, July 16, 2012 10:24 AM
To: 'Gail'
Cc: 'Yi-Fang Lee'; 'Hung, Sonya'
Subject: RE: 2008 Evaluation and Program Planning article: Practices and challenges in educational program evaluation in the Asia-Pacific region: Results of a Delphi Study

Gail, we are pleased to grant you permission and all that we ask is that you cite us in your references. We wish the best of luck on your research and if you are at AEA, please say hello. Best, JWA

From: Gail
Sent: Saturday, July 14, 2012 5:54 PM
To: Altschuld, James
Subject: 2008 Evaluation and Program Planning article: Practices and challenges in educational program evaluation in the Asia-Pacific region: Results of a Delphi Study

Dr. Altschuld,

By way of introduction, I am a doctoral candidate at the University of North Florida. My dissertation study is focused on evaluation capacity in the nonprofit sectors. As such, I will be conducting a Delphi Study with funders and providers regarding the use of program evaluation in nonprofit organizations. I found the article (see above) very informative and helpful in helping me design my study. I would like to use the some or all of the five open-ended questions in your article for my study. I would modify the wording slightly to put the focus on nonprofits in my study setting instead of the use of "country" in your questions. Please let me know if this is acceptable to you and your colleagues.

Thank you for your consideration to this matter.

Peace,

Gail Patin

winmail.dat
7K

Gmail by Google

Gail Patin

FW: Request to use your questions

1 message

Gail
To: Gail

Sun, Apr 7, 2013 at 4:05 PM

———

From: Gail
Sent: Tuesday, July 17, 2012 11:28 PM
To: 'Carman, Joanne'
Subject: RE: Request to use your questions

Dr. Carman,

Thank you for allow me to use or adapt your questions. Also thank you for your feedback on the wording of the questions. Your comments absolutely make sense and will help me refine the questions better.

Peace,

Gail Patin

———

From: Carman, Joanne
Sent: Sunday, July 15, 2012 4:16 PM
To: Gail
Subject: RE: Request to use your questions

Hi Gail -

Great to hear from you. Glad to hear you are moving along with your research, and I am happy to see you replicate or adapt any of the questions I used.

I have looked at your questions. They seem fine to me -- but .... ;-)

If I had to make a prediction, someone will ask you what you mean by 'effective'.... so, you might want to give that some thought or consider rephrasing (i.e., helpful, useful?...).

Same thing for 'quality' -- as a reader, I wonder if quality will mean different things to different people (i.e., people who have training in research methods will think about quality in terms of methodological rigor, others who don't might think about quality in terms of ease of utility or use, etc.). Of course, this would be interesting in and of itself - but I am guessing this will come up in the interviews, so you might want to give this some thought ahead of time and think about how you will want to respond.

I look forward to seeing the results of your research!

Best regards,

....Joanne

Joanne G. Carman, Ph.D. | Associate Professor

UNC Charlotte | Dept. of Political Science & Public Administration

9201 University City Blvd. | Charlotte, NC 28223

Phone:

jgcarman

From: Gail
Sent: Saturday, July 14, 2012 7:34 PM
To: Carman, Joanne
Subject: Request to use your questions

Dr. Carman,

I am a doctoral candidate at the University of North Florida and I wrote to you at the end of last year (see below) regarding a study you conducted on evaluation capacity in nonprofits (Evaluation Capacity and Nonprofit Organizations: Is the glass half-empty or half-full?). You were so kind as to send me the questions you used on your surveys as well as other helpful information.

I am designing a Delphi study in the area of evaluation and evaluation capacity building as it relates to the nonprofit sector in Northeast Florida. Additionally, I am going to conduct follow-up interviews with some of the panelists. As such, I will use generative/exploratory questions for the first round of the Delphi survey and questions for the interview stage of the study. Below are questions from your survey that I used verbatim or adapted to my purposes:

* How can nonprofit organizations effectively use program evaluation information collected?

What are the resources needed to build capacity in order to sustain program evaluation efforts in the Northeast Florida nonprofit sector?

* What would you identify as the barriers faced by the nonprofit sector in Northeast Florida for implementing quality program evaluation?

I am writing to see if this is acceptable to you and also to thank you for the research you have done in this area. Your work is very thought provoking and has not only helped me as a student but as a Chief Operating Officer for the nonprofit agency I work in.

Thank you for your consideration to this matter.

Peace,

Gail Patin

Hi Gail,

Thank you so much for your email and interest in my research.

I am glad you found the article interesting. I am happy to share the survey with you - it is actually the second version of the survey- so I have attached the first version as well so you can see how the instrument evolved.

And, not to overwhelm with you with information - but I am attaching a third file - it is the article that, for the most part, reflects the theoretical framework I used when creating the survey.

If you find that you are interesting in replicating the survey, I would be happy to have a conversation with you about what worked well, what didn't, and ways it can probably be improved in the future.

Best regards,

Joanne

Joanne G. Carman, Ph.D. | Associate Professor

UNC Charlotte | Dept. of Political Science & Public Administration

9201 University City Blvd. | Charlotte, NC 28223

Phone

jgcarman

From: Gail
Sent: Thursday, December 15, 2011 8:13 PM
To: Carman, Joanne
Subject: Evaluation Capacity and Nonprofit Organizations: Is the glass half-empty or half-full?

Dr. Carman,

By way of introduction, I am a doctoral candidate at the University of North Florida in the Educational Leadership program. Additionally, I am the Chief Operating Officer for a large nonprofit in Jacksonville, Florida and have over 25 years experience in nonprofit with most of them in management. As such, my dissertation/study is focusing on program evaluation capacity. I found your article (listed above) very informative and provided me with a wealth of information as I begin my study. In your study, you used a survey instrument to collect data. Is it possible for me to have a copy of that instrument? I am not sure what direction my study will take as I am at the beginning of the process. However, I thought it would be helpful and interesting to review the actual instrument you used for this article/study.

I would appreciate any assistance you could provide.

Peace,

Gail Patin

---

winmail.dat
15K

Appendix F

Initial Second Round Delphi Survey

Default Question Block

Welcome to Round 2 of my dissertation study.

The first round of the survey generated over 300 responses. Your responses were invaluable, thoughtful, informative, and provided a wealth of data to analyze. I took the information from this first round and reduced it to 8 categories in each of the two sections (role of program evaluation and capacity for program evaluation).

This second round should take much less time as I am asking you to rank order the categories. First, however, you will need to acknowledge your consent to participate then you can proceed to the rank ordering of categories. I also included a few demographic questions at the end of this survey.

<u>This survey is estimated to take between 10 - 20 minutes.</u>

Here are some tips that may help you navigate this survey.

1. This survey may be completed in more than one session. In order to do so, please follow the steps listed below so the responses you put in will be saved.

Be sure to save the email I sent as it contains the link that will take you back to your survey. After you enter your responses for any given question, you must click the "Save/Next" button or the "Previous" button in order to save your responses. If you close out of the program before doing this, your responses will not be saved. After saving your responses, as noted above, you may simply close the web tab/page by clicking on the "X" button (usually in the top right hand corner of a web tab/or page) to leave the survey. Use the link in the email to return to your saved survey responses.

<u>2.</u> Do not click the "Finish" button until you are completely finished and ready to submit your responses.

3. If you have any questions/issues/concerns, do not hesitate to contact me (Gail Patin) at or You may contact me at anytime (even early mornings or later in the evening). I want to make this experience as easy as possible for you.

I appreciate your participate in this multi-round survey process. Your expertise is invaluable.

Dear Participant:

My name is Gail Patin and I am a student in the Educational Leadership (Ed.D.) doctoral program at the University of North Florida (UNF). I am conducting a research study on program evaluation efforts in the nonprofit sector. This study aims to learn the perceptions of leaders in funding organizations and human services agencies regarding the role of and capacity for program evaluation in the Northeast Florida nonprofit sector.

I invite you to take part in this study as you are identified as a leader expert in a nonprofit or funding organization. If you take part in my study, you will be asked to respond to an online multi-round Delphi survey regarding the role of and capacity for program evaluation in the nonprofit sector. The Delphi Survey consists of two to four (2-4) rounds. The first questionnaire is exploratory in nature where I, as the researcher, ask for your expert opinion on a series of broad questions. The first survey is estimated to take between 20-40 minutes. Subsequent surveys are expected to take half the amount of time as you will be rating and/or priority ranking information I condensed from the initial questionnaire. Data gathered from the Delphi survey will be confidential to the extent allowable by law. No one other than myself will know your identity and I will maintain your individual responses with the strictest confidentiality. I will not share your name, the name of your organization or other identifying information.

As a direct benefit for taking place in the Delphi survey, I will provide you with a final copy of the Delphi results. Additionally, others may benefit from the information we learn from the results of this study. However, you will not be compensated for your participation. There are no foreseeable risks for taking part in this study. Participation is voluntary and there are no penalties for skipping questions or withdrawing your participation. Thus, you may choose to withdraw from this study at any time with no penalty or loss of benefits you would otherwise be entitled to receive. Additionally, I may contact you for an interview following the Delphi survey of the interview. Participation in the interview is also voluntary.

If you have any questions or concerns about this study, please contact me or my professor, Dr. Katherine Kasten. If you have questions about your rights as a participant, you may contact the University of North Florida's Institutional Review Board Vice Chairperson, Dr. Krista Paulson, at (904) 620-1650 or kpaulsen@unf.edu. I thank you for your consideration.

Sincerely,

To print this informed consent page, look for a print icon at the top of the web browser or under the "File" drop down menu.

- ○ By clicking here, I confirm that I am at least 18 years of age and understand and accept the above information. I agree and consent to participate in this study.
- ○ No, I do not wish to participate in this study.

Below are 8 categories regarding the role of program evaluation in the northeast Florida nonprofit sector as indicated by your responses from the first survey. As expert consultants in this study, I am asking you to critically discriminate between these important categories by ranking from most important (1) to least important (8).

The role of program evaluation in the northeast Florida nonprofit sector is to:

	1	2	3	4	5	6	7	8
Validate Organizational Credibility through mission/vision alignment, accountability & transparency, justification of value/importance of organization, and strategic planning efforts.	○	○	○	○	○	○	○	○
Determine Resource Allocation such as a guide for use of resources, cost effectiveness measures, efficiency determination, budget & cost considerations, and asset management.	○	○	○	○	○	○	○	○
Inform Program Management Decisions such as programmatic needs assessment, program design/development, and implementation.	○	○	○	○	○	○	○	○
Facilitate Quality Assurance through confirmation of program expectations, verification of goals/objectives achievement, and to inform program improvement strategies that includes elimination/adjustments/corrections of negative/ineffective programs or practices.	○	○	○	○	○	○	○	○

	1	2	3	4	5	6	7	8
Assess Impact of Program such as determination of improvement/benefit in lives of program participants, program effectiveness/success, outcomes, measurement of change.	○	○	○	○	○	○	○	○
Advance Organizational Learning through reflective practice, increased knowledge, feedback opportunities, a focus for education/training, and direction for change.	○	○	○	○	○	○	○	○
Enhance Communication with Multiple Stakeholders (e.g. staff, public, program participants, other organizations, policy makers) through sharing of program results, marketing strategies, and reciprocal feedback opportunities with program participants.	○	○	○	○	○	○	○	○
Cultivate Funding Collaborations such as providing data regarding the efficacy of program(s), providing information learned from evaluations, providing information regarding program improvements/development, satisfying funder requirements, and a means for garnering financial support.	○	○	○	○	○	○	○	○

Below are the categories regarding the capacity for program evaluation in the northeast Florida nonprofit sector as indicated by your responses from the first survey. As expert consultants in this study, I am asking you to critically discriminate between these important categories by ranking from most important (1) to least important (8).

The capacity for program evaluation in the northeast Florida nonprofit sector relies on these essential elements:

	1	2	3	4	5	6	7	8
Sufficient Time available for staff to plan, implement, analyze, and reflect on results from program evaluation.	○	○	○	○	○	○	○	○
Sufficient Financial Resources specifically designated for program evaluation efforts.	○	○	○	○	○	○	○	○
Sufficient Human Resources such as skilled and designated evaluation staff or access to/partnerships with experts (professional evaluators, researcher community, program evaluation technical assistance consultants, etc…)	○	○	○	○	○	○	○	○
A Positive Culture (organizational and sector-wide) that advances program evaluation efforts such as prioritizing resources, willingness to accept feedback, openness to change, and a readiness to reflect and act results.	○	○	○	○	○	○	○	○
Realistic Expectations from the philanthropic community regarding program evaluation efforts such as evaluation design, expected outcomes, reporting requirements, and their response to the results.	○	○	○	○	○	○	○	○
Functional Program Evaluation Designs/Methods that take into account feasibility, appropriate measurement tools, access to data, infrastructure to accommodate data, and issues related to engaging program participants in the evaluative process.	○	○	○	○	○	○	○	○
Ongoing Collaboration in and between the philanthropic and provider communities through negotiating appropriate evaluation strategies, sharing resources, providing results, and opportunities for discussion.	○	○	○	○	○	○	○	○
Ongoing Training for providers and funders regarding program evaluation strategies/best practices and available resources.	○	○	○	○	○	○	○	○

Type of organization you represent

○ Philanthropic (funding) Organization

○ Nonprofit Provider Organization

Organizational Budget Size

- < $150,000
- $150,000 - $349,999
- $350,000 - $499,999
- $500,000 - $749,000
- $750,000 - $1 Million
- $1 Million+ to $5 Million
- $5 Million+ to $10 Million
- $10 Million+

How many employees work in your organization?

- 1-4
- 5-9
- 10-19
- 20-49
- 50-99
- 100+

How long have you worked in this field (philanthropic/funder and/or nonprofit provider)? Note: If you have worked both in provider organizations and philanthropic/funder organizations then count the cumulative amount of experience from these venues. Similarly, if you have worked in different provider organizations and/or philanthropic/funder organizations then also count all (cumulative) of the years.

- 1 year or less
- 2-4 years
- 5 Years
- 6-9 years
- 10+ years

What is the highest level of education you have completed?

- Less than High School
- High School / GED
- Some College
- 2-year College Degree
- 4-year College Degree
- Masters Degree
- Doctoral Degree
- Professional Degree (JD, MD)

What is your gender?

- Male
- Female

What is your race? (U.S. Census)

- White/Caucasian
- African American
- American Indian or Alaska Native
- Asian
- Native Hawaiian or Other Pacific Islander
- Other

Hispanic or Latino Origin?

- No
- Yes

What is your current age? (U.S. Census)

- Under 20
- 20 to 24
- 25 to 34
- 35 to 44
- 45 to 54
- 55 to 64
- 65 or over

Additional Comments (optional)

PLEASE STOP & READ

If you are **NOT FINISHED** with the survey and wish to come back later to complete, **DO NOT** click on the "Finish" button below.

Instead, simply close the web tab/page by clicking on the "X" button (usually in the top right hand corner of a web tab/or page) to leave the survey. Use the link in the email to return to your saved survey responses. This will save your responses for when you return to complete the survey.

If you have any questions/issues/concerns, do not hesitate to contact me (Gail Patin) at or

You may contact me at anytime (even early mornings or later in the evening). I want to make this experience as easy as possible for you.

If you have completed the survey, then click on the "Finish" button below.

Appendix G

Revised Second Round Delphi Survey

Default Question Block

Welcome to Round 2 of my dissertation study.

The first round of the survey generated over 300 responses. Your responses were invaluable, thoughtful, informative, and provided a wealth of data to analyze. I took the information from this first round and reduced it to 8 categories in each of the two sections (role of program evaluation and capacity for program evaluation).

This second round should take much less time as I am asking you to rate the categories. First, however, you will need to acknowledge your consent to participate then you can proceed to rating the categories. I also included a few demographic questions at the end of this survey.

This survey is estimated to take between 10 - 20 minutes.

Here are some tips that may help you navigate this survey.

1. This survey may be completed in more than one session. In order to do so, please follow the steps listed below so the responses you put in will be saved.

Be sure to save the email I sent as it contains the link that will take you back to your survey. After you enter your responses for any given question, you must click the "Save/Next" button or the "Previous" button in order to save your responses. If you close out of the program before doing this, your responses will not be saved. After saving your responses, as noted above, you may simply close the web tab/page by clicking on the "X" button (usually in the top right hand corner of a web tab/or page) to leave the survey. Use the link in the email to return to your saved survey responses.

2. Do not click the "Finish" button until you are completely finished and ready to submit your responses.

3. If you have any questions/issues/concerns, do not hesitate to contact me (Gail Patin) at . You may contact me at anytime (even early mornings or later in the evening). I want to make this experience as easy as possible for you.

I appreciate your participate in this multi-round survey process. Your expertise is invaluable.

Dear Participant:

My name is Gail Patin and I am a student in the Educational Leadership (Ed.D.) doctoral program at the University of North Florida (UNF). I am conducting a research study on program evaluation efforts in the nonprofit sector. This study aims to learn the perceptions of leaders in funding organizations and human services agencies regarding the role of and capacity for program evaluation in the Northeast Florida nonprofit sector.

I invite you to take part in this study as you are identified as a leader expert in a nonprofit or funding organization. If you take part in my study, you will be asked to respond to an online multi-round Delphi survey regarding the role of and capacity for program evaluation in the nonprofit sector. The Delphi Survey consists of two to four (2-4) rounds. The first questionnaire is exploratory in nature where I, as the researcher, ask for your expert opinion on a series of broad questions. The first survey is estimated to take between 20-40 minutes. Subsequent surveys are expected to take half the amount of time as you will be rating and/or priority ranking information I condensed from the initial questionnaire. Data gathered from the Delphi survey will be confidential to the extent allowable by law. No one other than myself will know your identity and I will maintain your individual responses with the strictest confidentiality. I will not share your name, the name of your organization or other identifying information.

As a direct benefit for taking place in the Delphi survey, I will provide you with a final copy of the Delphi results. Additionally, others may benefit from the information we learn from the results of this study. However, you will not be compensated for your participation. There are no foreseeable risks for taking part in this study. Participation is voluntary and there are no penalties for skipping questions or withdrawing your participation. Thus, you may choose to withdraw from this study at any time with no penalty or loss of benefits you would otherwise be entitled to receive. Additionally, I may contact you for an interview following the Delphi survey of the interview. Participation in the interview is also voluntary.

If you have any questions or concerns about this study, please contact me or my professor, Dr. Katherine Kasten. If you have questions about your rights as a participant, you may contact the University of North Florida's Institutional Review Board Vice Chairperson, Dr. Krista Paulson, at (904) 620-1650 or kpaulsen@unf.edu. I thank you for your consideration.

Sincerely,

To print this informed consent page, look for a print icon at the top of the web browser or under the "File" drop down menu.

- By clicking here, I confirm that I am at least 18 years of age and understand and accept the above information. I agree and consent to participate in this study.
- No, I do not wish to participate in this study.

Below are 8 categories created from your responses from the first survey regarding **the role of program evaluation in the nonprofit sector.**

As expert consultants in this study, I am asking you to critically rate the level of importance of each of these proposed categories.

The role of program evaluation in the northeast Florida nonprofit sector is to:

	1=Not At All Important	2=Not Very Important	3=Neither Important nor Unimportant	4=Important	5=Very Important
Validate Organizational Credibility through mission/vision alignment, accountability & transparency, justification of value/importance of organization, and strategic planning efforts.	○	○	○	○	○
Determine Resource Allocation such as a guide for use of resources, cost effectiveness measures, efficiency determination, budget & cost considerations, and asset management.	○	○	○	○	○
Inform Program Management Decisions such as programmatic needs assessment, program design/development, and implementation.	○	○	○	○	○
Facilitate Quality Assurance through confirmation of program expectations, verification of goals/objectives achievement, and to inform program improvement strategies that includes elimination/adjustments /corrections of negative/ineffective programs or practices.	○	○	○	○	○
Assess Impact of Program such as determination of improvement/benefit in lives of program participants, program effectiveness/success, outcomes, measurement of change.	○	○	○	○	○
Advance Organizational Learning through reflective practice, increased knowledge, feedback opportunities, a focus for education/training, and direction for change.	○	○	○	○	○
Enhance Communication with Multiple Stakeholders (e.g. staff, public, program participants, other organizations, policy makers) through sharing of program results, marketing strategies, and reciprocal feedback opportunities with program participants.	○	○	○	○	○
Cultivate Funding Collaborations such as providing data regarding the efficacy of program(s), providing information learned from evaluations, providing information regarding program improvements/development, satisfying funding requirements, and a means for garnering financial support.	○	○	○	○	○

Below are the categories created from your responses from the first survey regarding the **capacity for program evaluation** in the northeast Florida nonprofit sector.

As expert consultants in this study, I am asking you to critically rate your agreement regarding each of these proposed categories.

The capacity for program evaluation in the northeast Florida nonprofit sector relies on these essential elements:

	1=Strongly Disagree	2=Disagree	3=Neither Agree nor Disagree	4=Agree	5=Strongly Agree
Sufficient Time available for staff to plan, implement, analyze, and reflect on results from program evaluation.	○	○	○	○	○
Sufficient Financial Resources specifically designated for program evaluation efforts.	○	○	○	○	○
Sufficient Human Resources such as skilled and designated evaluation staff or access to/partnerships with experts (professional evaluators, researcher community, program evaluation technical assistance consultants, etc…)	○	○	○	○	○
A Positive Culture (organizational and sector-wide) that advances program evaluation efforts such as prioritizing resources, willingness to accept feedback, openness to change, and a readiness to reflect and act results.	○	○	○	○	○
Realistic Expectations from the philanthropic community regarding program evaluation efforts such as evaluation design, expected outcomes, reporting requirements, and their response to the results.	○	○	○	○	○
Functional Program Evaluation Designs/Methods that take into account feasibility, appropriate measurement tools, access to data, infrastructure to accommodate data, and issues related to engaging program participants in the evaluative process.	○	○	○	○	○
Ongoing Collaboration in and between the philanthropic and provider communities through negotiating appropriate evaluation strategies, sharing resources, providing results, and opportunities for discussion.	○	○	○	○	○
Ongoing Training for provider and funding personnel regarding program evaluation strategies/best practices and available resources.	○	○	○	○	○

Type of organization you represent

- ○ Philanthropic (funding) Organization
- ○ Nonprofit Provider Organization

Organizational Budget Size

- ○ < $150,000
- ○ $150,000 - $349,999
- ○ $350,000 - $499,999
- ○ $500,000 - $749,000
- ○ $750,000 - $1 Million
- ○ $1 Million+ to $5 Million
- ○ $5 Million+ to $10 Million
- ○ $10 Million+

How many employees work in your organization?

- ○ 1-4
- ○ 5-9
- ○ 10-19
- ○ 20-49
- ○ 50-99
- ○ 100+

How long have you worked in this field (philanthropic/funding and/or nonprofit provider)? Note: If you have worked both in provider organizations and philanthropic/funding organizations then count the cumulative amount of experience from these venues. Similarly, if you have worked in different provider organizations and/or philanthropic/funding organizations then also count all (cumulative) of the years.

- 1 year or less
- 2-4 years
- 5 Years
- 6-9 years
- 10+ years

What is the highest level of education you have completed?

- Less than High School
- High School / GED
- Some College
- 2-year College Degree
- 4-year College Degree
- Masters Degree
- Doctoral Degree
- Professional Degree (JD, MD)

What is your gender?

- Male
- Female

What is your race? (U.S. Census)

- White/Caucasian
- African American
- American Indian or Alaska Native
- Asian
- Native Hawaiian or Other Pacific Islander
- Other

Hispanic or Latino Origin?

- No
- Yes

What is your current age? (U.S. Census)

- Under 20
- 20 to 24
- 25 to 34
- 35 to 44
- 45 to 54
- 55 to 64
- 65 or over

Additional Comments (optional)

PLEASE STOP & READ

If you are **NOT FINISHED** with the survey and wish to come back later to complete, **DO NOT** click on the "Finish" button below.

Instead, simply close the web tab/page by clicking on the "X" button (usually in the top right hand corner of a web tab/or page) to leave the survey. Use the link in the email to return to your saved survey responses. This will save your responses for when you return to complete the survey.

If you have any questions/issues/concerns, do not hesitate to contact me (Gail Patin) at

You may contact me at anytime (even early mornings or later in the evening). I want to make this experience as easy as possible for you.

If you have completed the survey, then click on the "Finish" button below.

Appendix H

Third Round Delphi Survey

Default Question Block

Welcome to the 3rd Round of the Delphi Survey. As mentioned previously, Delphi studies usually consist of 2-4 rounds.

The second survey round produced results that indicated agreement of six essential elements of program evaluation **capacity** for the northeast Florida nonprofit sector.

This follow-up round asks you to rate to what **degree these essential program capacity elements are currently present** in the Northeast Florida nonprofit sector as a whole.

<u>This survey is estimated to take between 5-10 minutes.</u>

Here are some tips that may help you navigate this survey.

1. This survey may be completed in more than one session. In order to do so, please follow the steps listed below so the responses you put in will be saved.

Be sure to save the email I sent as it contains the link that will take you back to your survey. After you enter your responses for any given question, you must click the "Save/Next" button or the "Previous" button in order to save your responses. If you close out of the program before doing this, your responses will not be saved. After saving your responses, as noted above, you may simply close the web tab/page by clicking on the "X" button (usually in the top right hand corner of a web tab/or page) to leave the survey. Use the link in the email to return to your saved survey responses.

2. Do not click the "Finish" button until you are completely finished and ready to submit your responses.

3. If you have any questions/issues/concerns, do not hesitate to contact me (Gail Patin) at or You may contact me at anytime (even early mornings or later in the evening). I want to make this experience as easy as possible for you.

I appreciate your participate in this multi-round survey process. Your expertise is invaluable.

Dear Participant:

My name is Gail Patin and I am a student in the Educational Leadership (Ed.D.) doctoral program at the University of North Florida (UNF). I am conducting a research study on program evaluation efforts in the nonprofit sector. This study aims to learn the perceptions of leaders in funding organizations and human services agencies regarding the role of and capacity for program evaluation in the Northeast Florida nonprofit sector.

I invite you to take part in this study as you are identified as a leader expert in a nonprofit or funding organization. If you take part in my study, you will be asked to respond to an online multi-round Delphi survey regarding the role of and capacity for program evaluation in the nonprofit sector. The Delphi Survey consists of two to four (2-4) rounds. The first questionnaire is exploratory in nature where I, as the researcher, ask for your expert opinion on a series of broad questions. The first survey is estimated to take between 20-40 minutes. Subsequent surveys are expected to take half the amount of time as you will be rating and/or priority ranking information I condensed from the initial questionnaire. Data gathered from the Delphi survey will be confidential to the extent allowable by law. No one other than myself will know your identity and I will maintain your individual responses with the strictest confidentiality. I will not share your name, the name of your organization or other identifying information.

As a direct benefit for taking place in the Delphi survey, I will provide you with a final copy of the Delphi results. Additionally, others may benefit from the information we learn from the results of this study. However, you will not be compensated for your participation. There are no foreseeable risks for taking part in this study. Participation is voluntary and there are no penalties for skipping questions or withdrawing your participation. Thus, you may choose to withdraw from this study at any time with no penalty or loss of benefits you would otherwise be entitled to receive. Additionally, I may contact you for an interview following the Delphi survey of the interview. Participation in the interview is also voluntary.

If you have any questions or concerns about this study, please contact me or my professor, Dr. Katherine Kasten. If you have questions about your rights as a participant, you may contact the University of North Florida's Institutional Review Board Vice Chairperson, Dr. Krista Paulson, at . I thank you for your consideration.

Sincerely,

To print this informed consent page, look for a print icon at the top of the web browser or under the "File" drop down menu.

- ○ By clicking here, I confirm that I am at least 18 years of age and understand and accept the above information. I agree and consent to participate in this study.
- ○ No, I do not wish to participate in this study.

Below are the essential program evaluation capacity elements agreed upon from your responses from the second survey as determined by a mean/average score of at least 4. The elements are listed in the order of highest mean/average to lowest mean/average.

As expert consultants in this study, I am asking you to critically rate to **what degree each of these program evaluation capacity elements currently are present** in the northeast Florida nonprofit sector as a whole.

*The **degree to which each of these essential elements of capacity for program evaluation are present*** in the northeast Florida nonprofit sector is:

	1=Usually Not Present	2=Often Not Present	3=Sometimes Present	4=Often Present	5=Generally Present
Sufficient Time available for staff to plan, implement, analyze, and reflect on results from program evaluation. Mean=4.50	○	○	○	○	○
Functional Program Evaluation Designs/Methods that take into account feasibility, appropriate measurement tools, access to data, infrastructure to accommodate data, and issues related to engaging program participants in the evaluative process. Mean=4.33	○	○	○	○	○
Sufficient Human Resources such as skilled and designated evaluation staff or access to/partnerships with experts (professional evaluators, researcher community, program evaluation technical assistance consultants, etc…) Mean=4.25	○	○	○	○	○
A Positive Culture (organizational and sector-wide) that advances program evaluation efforts such as prioritizing resources, willingness to accept feedback, openness to change, and a readiness to reflect and act results. Mean=4.17	○	○	○	○	○
Sufficient Financial Resources specifically designated for program evaluation efforts. Mean=4.17	○	○	○	○	○
Ongoing Collaboration in and between the philanthropic and provider communities through negotiating appropriate evaluation strategies, sharing resources, providing results, and opportunities for discussion. Mean=4.00	○	○	○	○	○

Please feel free to write any additional comments or observations (optional).

PLEASE STOP & READ

If you are **NOT FINISHED** with the survey and wish to come back later to complete, **DO NOT** click on the "Finish" button below.

Instead, simply close the web tab/page by clicking on the "X" button (usually in the top right hand corner of a web tab/or page) to leave the survey. Use the link in the email to return to your saved survey responses. This will save your responses for when you return to complete the survey.

If you have any questions/issues/concerns, do not hesitate to contact me (Gail Patin) at

You may contact me at anytime (even early mornings or later in the evening). I want to make this experience as easy as possible for you.

If you have completed the survey, then click on the "Finish" button below.

Block 1

Appendix I

Example Fourth Round Delphi Survey

Default Question Block

Dear Expert Consultant,

Here are the results of **your** individual, group (funder or provider), and combined results from the multi-round Delphi survey. Please note that this is an individualized result because it has your own rating. As such, I had to create 12 different surveys for each individual expert consultant.

A key component of the Delphi survey is that you, as an expert consultant in the study, get to review the results in order to provide comments (to your colleagues involved in this study in an anonymous way) supporting your ratings, overall comments on the findings, and it is a last opportunity for you to change your rating in any given category. Of course, you do not have to change your rating if you feel it still reflects your perception of the subject.

Please take some time to analyze the results and compare your answer to the groups' combined results. **This is an opportunity to change your response for each category and/or provide a brief explanation of your rating.** This is important as it allows you one last opportunity to consider your answer based upon your responses as compared to the responses of your colleagues.

Since this final round will take some reflection on your part, the estimated time is between 30-45 minutes.

Here are some tips that may help you navigate this survey.

1. This survey may be completed in more than one session. In order to do so, please follow the steps listed below so the responses you put in will be saved.

Be sure to save the email I sent as it contains the link that will take you back to your survey. After you enter your responses for any given question, you must click the "Save/Next" button or the "Previous" button in order to save your responses. If you close out of the program before doing this, your responses will not be saved. After saving your responses, as noted above, you may simply close the web tab/page by clicking on the "X" button (usually in the top right hand corner of a web tab/or page) to leave the survey. Use the link in the email to return to your saved survey responses.

2. Do not click the "Finish" button until you are completely finished and ready to submit your responses.

3. If you have any questions/issues/concerns, do not hesitate to contact me (Gail Patin) at You may contact me at anytime (even early mornings or later in the evening). I want to make this experience as easy as possible for you.

I appreciate your participate in this multi-round survey process. Your expertise is invaluable.

Dear Participant:

My name is Gail Patin and I am a student in the Educational Leadership (Ed.D.) doctoral program at the University of North Florida (UNF). I am conducting a research study on program evaluation efforts in the nonprofit sector. This study aims to learn the perceptions of leaders in funding organizations and human services agencies regarding the role of and capacity for program evaluation in the Northeast Florida nonprofit sector.

I invite you to take part in this study as you are identified as a leader expert in a nonprofit or funding organization. If you take part in my study, you will be asked to respond to an online multi-round Delphi survey regarding the role of and capacity for program evaluation in the nonprofit sector. The Delphi Survey consists of two to four (2-4) rounds. The first questionnaire is exploratory in nature where I, as the researcher, ask for your expert opinion on a series of broad questions. The first survey is estimated to take between 20-40 minutes. Subsequent surveys are expected to take half the amount of time as you will be rating and/or priority ranking information I condensed from the initial questionnaire. Data gathered from the Delphi survey will be confidential to the extent allowable by law. No one other than myself will know your identity and I will maintain your individual responses with the strictest confidentiality. I will not share your name, the name of your organization or other identifying information.

As a direct benefit for taking place in the Delphi survey, I will provide you with a final copy of the Delphi results. Additionally, others may benefit from the information we learn from the results of this study. However, you will not be compensated for your participation. There are no foreseeable risks for taking part in this study. Participation is voluntary and there are no penalties for skipping questions or withdrawing your participation. Thus, you may choose to withdraw from this study at any time with no penalty or loss of benefits you would otherwise be entitled to receive. Additionally, I may contact you for an interview following the Delphi survey of the interview. Participation in the interview is also voluntary.

If you have any questions or concerns about this study, please contact me or my professor, Dr. Katherine Kasten. If you have questions about your rights as a participant, you may contact the University of North Florida's Institutional Review Board Vice Chairperson, Dr. Krista Paulson, I thank you for your consideration.

Sincerely,

To print this informed consent page, look for a print icon at the top of the web browser or under the "File" drop down menu.

- ○ By clicking here, I confirm that I am at least 18 years of age and understand and accept the above information. I agree and consent to participate in this study.
- ○ No, I do not wish to participate in this study.

Below are 8 categories (ranked from highest average to lowest average) created from your responses from the second survey regarding **the role of program evaluation in the nonprofit sector.**

For each category, I included descriptive statistics of the results grouped by each different panel of experts (funders & providers) and then combined both groups. In the upper right hand corner of the table, your individual rating is also noted.

Under the table, I may have a comment if your rating for any particular category is outside the middle range of responses or if someone provided comments from the 2nd or 3rd rounds. I would ask that you consider your original rating in context of your colleagues' responses. **You may change your rating and/or provide feedback for your answer.** Of course, you do not have to change your rating if you feel it still reflects your perception of the subject.

If you wish to change your rating please put it in the text box to the right of the provided information. Additionally, comments may go in that box as well.

If you do not wish to change your rating or put any comments please put "**NC**" in the box.

As a reminder, below is the rating scaled used and the statement for this category was: *The role of program evaluation in the northeast Florida nonprofit sector is:*				
1	2	3	4	5
Not At All Important	Not Very Important	Neither Important nor Unimportant	Important	Very Important

1. Assess Impact of Program such as determination of improvement/benefit in lives of program participants, program effectiveness/success, outcomes, measurement of change.

Your Response: 5	Funders	Providers	Combined
Average	5	4.71	4.83
Minimum	5	4	4
Maximum	5	5	5

Comments:

2. Advance Organizational Learning through reflective practice, increased knowledge, feedback opportunities, a focus for education/training, and direction for change.

Your Response: 4	Funders	Providers	Combined
Average	4.80	4.29	4.5
Minimum	4	3	3
Maximum	5	5	5

Comments: Your rating was lower than other "Funders" ratings.

3. Cultivate Funding Collaborations such as providing data regarding the efficacy of program(s), providing information learned from evaluations, providing information regarding program improvements/development, satisfying funding requirements, and a means for garnering financial support.

Your Response: 4	Funders	Providers	Combined
Average	4.00	4.71	4.42
Minimum	4	4	4
Maximum	4	5	5

Comments:

Below are 8 categories (ranked from highest average to lowest average) created from your responses from the second survey regarding **the capacity for program evaluation in the nonprofit sector.**

For each category, I included descriptive statistics of the results grouped by each different panel of experts (funders & providers) and then combined both groups. In the upper right hand corner of the table, your individual rating is also noted.

Under the table, I may have a comment if your rating for any particular category is outside the middle range of responses or if someone provided comments from the 2nd or 3rd rounds. I would ask that you consider your original rating in context of your colleagues' responses. **You may change your rating and/or provide feedback for your answer.** Of course, you do not have to change your rating if you feel it still reflects your perception of the subject.

If you wish to change your rating please put it in the text box to the right of the provided information. Additionally, comments may go in that box as well.

If you do not wish to change your rating or put any comments please put "**NC**" in the box.

As a reminder, below is the rating scaled used and the statement for this category was ***The capacity for program evaluation in the northeast Florida nonprofit sector relies on these essential elementss:***

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

1. **Sufficient Time** available for staff to plan, implement, analyze, and reflect on results from program evaluation.

Your Response: 5	Funders	Providers	Combined
Average	4.60	4.43	4.5
Minimum	3	3	3
Maximum	5	5	5

Comments:

2. **Functional Program Evaluation Designs/Methods** that take into account feasibility, appropriate measurement tools, access to data, infrastructure to accommodate data, and issues related to engaging program participants in the evaluative process.

Your Response: 4	Funders	Providers	Combined
Average	4.2	4.43	4.33
Minimum	3	2	2
Maximum	5	5	5

Comments:

3. **Sufficient Human Resources** such as skilled and designated evaluation staff or access to/partnerships with experts (professional evaluators, researcher community, program evaluation technical assistance consultants, etc...)

Your Response:	Funders	Providers	Combined

Reality Check:

Below are 6 categories (ranked from highest average to lowest average) created from your responses from the third survey regarding **the degree to which each of these essential elements of capacity for program evaluation are present in the northeast Florida nonprofit sector.**

For each category, I included descriptive statistics of the results grouped by each different panel of experts (funders & providers) and then combined both groups. In the upper right hand corner of the table, your individual rating is also noted.

Under the table, I may have a comment if your rating for any particular category is outside the middle range of responses or if someone provided comments from the 2nd or 3rd rounds. I would ask that you consider your original rating in context of your colleagues' responses. **You may change your rating and/or provide feedback for your answer.** Of course, you do not have to change your rating if you feel it still reflects your perception of the subject.

If you wish to change your rating please put it in the text box to the right of the provided information. Additionally, comments may go in that box as well.

If you do not wish to change your rating or put any comments please put "**NC**" in the box.

As a reminder, below is the rating scaled used and the statement for this category was ***The degree to which each of these essential elements of capacity for program evaluation are present in the northeast Florida nonprofit sector is:***				
1	2	3	4	5
Usually Not Present	Often Not Present	Sometimes Present	Often Present	Generally Present

1. **A Positive Culture** (organizational and sector-wide) that advances program evaluation efforts such as prioritizing resources, willingness to accept feedback, openness to change, and a readiness to reflect and act results.

Your Response: 4	Funders	Providers	Combined
Average	3.6	3.29	3.42
Minimum	3	2	2
Maximum	5	5	5

Comments:

2. **Functional Program Evaluation Designs/Methods** that take into account feasibility, appropriate measurement tools, access to data, infrastructure to accommodate data, and issues related to engaging program participants in the evaluative process.

Your Response: 3	Funders	Providers	Combined
Average	3	2.86	2.92
Minimum	2	1	1
Maximum	4	5	5

Comments:

Optional additional comments regarding the role of and capacity for program evaluation in the nonprofit sector, overall results, participation in this Delphi study, etc...

PLEASE STOP & READ

If you are **NOT FINISHED** with the survey and wish to come back later to complete, **DO NOT** click on the "Finish" button below.

Instead, simply close the web tab/page by clicking on the "X" button (usually in the top right hand corner of a web tab/or page) to leave the survey. Use the link in the email to return to your saved survey responses. This will save your responses for when you return to complete the survey.

If you have any questions/issues/concerns, do not hesitate to contact me (Gail Patin) at

You may contact me at anytime (even early mornings or later in the evening). I want to make this experience as easy as possible for you.

If you have completed the survey, then click on the "Finish" button below.

Block 1

Appendix J

Transcriber Confidentiality Statement

Confidentiality Agreement
Transcription Services

In signing below, you are agreeing to respect the participant's right to privacy and that of the people and organizations that may be included in the information collected. You are required to respect people's right to confidentiality by not discussing the information collected in public, with friends or family members.

I, ________________________, transcriptionist, agree to maintain full confidentiality in regards to any and all audiotapes and documentation received from [Researcher Name] related to her doctoral study on [Title of Study]

Furthermore, I agree and understand:

1. To respect the participants' rights to privacy and that of the people and organizations that may be included in the information;
2. Not to discuss the information collected in public, with friends or family members;
3. I understand the importance of providing anonymity (if relevant) and confidentiality to research participants;
4. I understand that the research information may contain references to individuals or organizations in the community, other than the participant. I understand that this information is to be kept confidential;
5. When transcribing, I will be the only one to hear the tapes;
6. To hold in strictest confidence the identification of any individual that may be inadvertently revealed during the transcription of audio-taped interviews, or in any associated documents;
7. To not make copies of any audiotapes or computerized files of the transcribed interview texts, unless specifically requested to do so by [Researcher Name];
8. To store all study-related audiotapes and materials (electronic files, transcripts, etc…) in a safe, secure location at all times (e.g., not left unattended) as long as they are in my possession;
9. To return all audiotapes and study-related documents to [Researcher Name] in a complete and timely manner;
10. To permanently delete all electronic files containing study-related documents from my computer hard drive and any backup devices (if applicable);
11. To return all audio recordings and transcribed documents to [Researcher Name] once completed.

I am aware that I can be held legally liable for any breach of this confidentiality agreement, and for any harm incurred by individuals if I disclose identifiable information contained in the audiotapes and/or files to which I will have access.

Transcriber's name (printed):____
Transcriber's signature:____
Date:____
Witness Name (printed):____
Witness Signature____
Date____

Appendix K

IRB Approval

UNF
UNIVERSITY of NORTH FLORIDA.
Office of Research and Sponsored Programs
1 UNF Drive
Jacksonville, FL 32224-2665
904-620-2455 FAX 904-620-2457
Equal Opportunity/Equal Access/Affirmative Action Institution

MEMORANDUM

DATE: October 26, 2012

TO: Ms. Gail Patin

VIA: Dr. Katherine Kasten
LSCSM

FROM: Dr. Krista Paulsen, Vice Chairperson
On behalf of the UNF Institutional Review Board

RE: Review of New Project Revisions by the UNF Institutional Review Board IRB#378229-2: "The role of program evaluation in the nonprofit sector: An exploratory study of leaders' perceptions"

UNF IRB Number: 378229-2
Approval Date: 10-26-2012
Expiration Date: Exempt - None
Processed on behalf of UNF's IRB KLC

This is to advise you that your project, "The role of program evaluation in the nonprofit sector: An exploratory study of leaders' perceptions" was reviewed on behalf of the UNF Institutional Review Board and has been declared Exempt, Categories 2 & 4." Therefore, this project requires no further IRB oversight unless substantive changes are made.

This approval applies to your project in the form and content as submitted to the IRB for review. All participants must receive a stamped and dated copy of the approved informed consent document. Any variations or modifications to the approved protocol and/or informed consent forms that might increase risk to human participants must be submitted to the IRB prior to implementing the changes. Please see the UNF Standard Operating Procedures for additional information about what types of changes might elevate risk to human participants. Any unanticipated problems involving risk and any occurrence of serious harm to subjects and others shall be reported promptly to the IRB within 3 business days.

Your study has been approved as of 10/26/2012. Because your project was approved as exempt, no further IRB oversight is required for this project unless you intend to make a change that might elevate risk to participants. As an exempt study, continuing review will be unnecessary. When you are ready to close your project, please complete a Closing Report Form which can also be found in the documents library called "Forms and Templates" in IRBNet.

As you may know, **CITI Course Completion Reports are valid for 3 years**. Your completion report is valid through 9/22/2014 and Dr. Kasten's completion report is valid through 3/30/2014. If your completion report expires within the next 60 days or has expired, please take CITI's refresher course and contact us to let us know you have completed that training. If you have not yet completed your CITI training or if you need to complete the refresher course, please do so by following this link: http://www.citiprogram.org/. Should you have questions regarding your project or any other IRB issues, please contact the research integrity unit of the Office of Research and Sponsored Programs by emailing IRB@unf.edu or calling (904) 620-2455.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within UNF's records. All records shall be accessible for inspection and copying by authorized representatives of the department or agency at reasonable times and in a reasonable manner. A copy of this approval may also be sent to the dean and/or chair of your department.

UNF IRB Number: 378229-2
Approval Date: 10-26-2012
Expiration Date: Exempt - None
Processed on behalf of UNF's IRB KLC

Appendix L

Words/Ideas Clusters for the Role of Program Evaluation

The following eight clusters of words and ideas came from the Delphi panelists responses from the first exploratory survey round. The responses were generated from four questions related to the role of program evaluation:

- Q1: When you hear or read the phrase, program evaluation, what does this mean to you?
- Q2: What are the purposes for conducting program evaluations in nonprofit organizations?
- Q3: What are the benefits to nonprofit organizations for conducting program evaluations?
- Q5: How can nonprofit organizations use program evaluation information collected?

The Role of Program Evaluation: Word/Ideas Cluster #1

Labeled and defined as validate organizational credibility through mission/vision alignment, accountability & transparency, justification of value/importance of organization, and strategic planning efforts:

- Are overall strategies linked to vision & mission of organization
- Guiding principles: is the vision and mission of the organization clear to staff and community
- Helps with clarity in defining their mission, purpose, goals & activities and that they all align
- Help in understanding how to set/refine organization’s strategic mission
- Are services linked to target population
- Are services well defined
- Provides an accountability measure

- Provides transparency
- Clearly defines the importance of the organization
- Gives credibility when independent evaluation finds program successful
- Helps justify their (nonprofit) value
- Help with understanding organizational value
- Gives the organization the idea of how well we are doing and gives us a direction to head in
- Improves organization as a whole
- Give organization the ability to develop and refine operational plans if needed

The Role of Program Evaluation: Word/Ideas Cluster #2

Labeled and defined as determine resource allocation such as a guide for use of resources, cost effectiveness measures, efficiency determination, budget & cost considerations, and asset management:

- Helps nonprofits ensure their extremely limited resources are being used in a way that has the most benefit to the most people
- Determine best use of resources
- Help guide use of resources
- Determine allocation of agency resources
- Learn what is working and what is not to make better use of resources
- Determine if funding is needed and appropriate
- To determine financial reasonableness to continue program
- Determination of program funding
- Ensure effective use of resources

- Is program cost effective
- To determine cost effectiveness
- Help create efficiencies
- Resources spent efficiently
- Efficiency (2)
- Determine programs that are financially strong
- Cost
- Keep on track with budget
- Realign organizational assets (staffing) to enhance good results

The Role of Program Evaluation: Word/Ideas Cluster #3

Labeled and defined as inform program management decisions such as programmatic needs assessment, program design/development, and implementation:

- Assess need for program
- Assessment of need for program
- Does the previous defined need for program still exist
- You can evaluate a program at many different stages to determine the need for the program, how it is to be implemented and/or outcome or impact
- To assess strengths and weaknesses of program
- To determine continued need of program
- To determine if program is needed
- Determine most needed and effective programs to operate
- Produce data to reinforce program design

- Program design/management: determining capacity needs, directing resources, based on utilization
- Determine program design
- Program evaluation depends upon clear definition of intended outcomes & determination of the most appropriate means of assessing progress towards those outcomes
- Identify effective personnel and program designs
- What does the program entail?
- What results are anticipated?
- How will results be measured?
- A report to staff to generate new or enhanced program design
- Create evidence-based practices
- To help inform future plans of work
- Understanding/determining most effective possible approach to achieving outcome
- Determine how to implement

The Role of Program Evaluation: Word/Ideas Cluster #4

Labeled and defined as facilitate quality assurance through confirmation of program expectations, verification of goals/objectives achievement and to inform program improvement strategies that includes elimination/adjustments/corrections of negative/ineffective programs or practices:

- Program doing what is expected?
- Make sure program is doing what is expected
- Determine if effects are having the effect hoped and intended them to have
- To ensure program is doing what is expected

- To know if services are helping as intended
- Is the program doing what was intended
- Empirically prove the program does what it intends to do
- Understanding what results are being hoped for and if they are being achieved
- The question to be answered is did the program have the intended effect, if not, can it be improved and is it overall worthwhile
- Help to know if on right track with goals and objectives of programs/projects
- Everyone knows what the activities of goals are to achieve goals
- To see clear understanding of program goal(s) and know what they are
- Are the previously identified goals of the program being met
- Measurable results of program
- Revise intervention strategies to improve service delivery
- Shows where improvements, clarity, better efficiencies need to be made
- Improve programs (x2)
- Improvements in services
- Strengthen areas that need improvement
- Help produce a “better product” or service
- Which activities to continue and build upon
- Get better at providing services
- Capacity building: in order to recognize the areas which are the strongest and those that need improvement
- Process for quality improvement
- Learning efforts in areas that need to be increased or changes

- Improve program quality
- Program improvement
- Improved program
- Which activities need to be changed in order to improve the program
- To inform training for staff & volunteers on how to better facilitate the program
- Help with program improvement
- Provide information for corrective actions with programs.
- Identify ineffective practices
- Help prevent mistakes
- Determine what corrections need to be made if the program has unintended outcomes
- Tells when wrong
- If program is not doing what is expected, is something beneficial being done
- To eliminate programs outsources, better done by someone else, or that don't align with mission
- Decisions about whether to continue existing efforts or undertake a different set of efforts.
- Fix areas that are measuring as weaknesses
- Strengthen or eliminate ineffective weak programs
- Elimination if nothing positive results from program

The Role of Program Evaluation: Words/Ideas Cluster #5

Labeled and defined as assess impact of program such as determination of improvement/benefit in lives of program participants, program effectiveness/success, outcomes, measurement of change:

- Are goals making an impact and not just measuring outcomes
- To help understand the impacts of the program/project on the people we serve
- Assess program outcome/impact
- Measure impact (x2)
- Making improvements in lives of program participants
- Knowing efforts are reaping results
- Knowing whether anyone benefits from program and how he/she benefits
- Learning if activities are making a difference
- To determine if program is working
- 3rd party review of validity and effectiveness
- Evidence of program effectiveness
- Demonstrates effectiveness
- Effectiveness of program (x4)
- Measuring results, determining effectiveness of program
- An assessment of the effectiveness of a particular program in achieving its intended outcomes
- Determine program success
- Ensure program success
- Determining if interventions are working
- Measures success in a realistic way
- To prove program works or doesn’t (x2)
- Measurement of program outcomes Outcome measures: what changed as a result of the program

- Outcome measurements, the ability to measure changes in organizations
- Outcomes
- Whether sustainable change over the long-term has been created
- To track progress or lack of progress from year to year

The Role of Program Evaluation: Words/Ideas Cluster #6

Labeled and defined as advance organizational learning through reflective practice, increased knowledge, feedback opportunities, a focus for education/training, and direction for change:

- Organizational learning: an opportunity to determine if we are accomplishing the mission
- Organizational learning- what is working, what isn't and maybe why it isn't
- What are the quantifiers- i.e. what uniquely or specifically makes this successful (or is inhibiting it from being successful)- what should be sustained and replicated (if applicable), and what should be changed or eliminated?
- What has been learned from the program thus far that prompts changes to original thoughts
- Good program evaluation is at the heart of reflective practice. Without it, cannot have the periodic "gut check" and recalibration to keep an organization learning, evolving, and thriving.
- The grantees we work with who find the most value in their program evaluations are those that set aside time for leadership-executives and board members as well as staff-to thoroughly understand the evaluation and engage in reflection about what should be done based on the knowledge gleaned from the evaluation (keep doing the same, do more, do less, tweak, or do something different entirely). To me this seems like the first and most valuable use through it is often hard to make time for this deep and thorough reflection.

- Does the program work? What makes it work better or worse than other similar programs?
- Understanding what the program is doing
- Identify trends
- Identify possible solutions
- Identify growth opportunities
- Comparison with other programs attempting to achieve same outcome
- Helpful to discover positives and negative side effects of program process
- Help understand what public policies should change and how
- Feedback for ongoing education for staff
- Feedback for staff delivering the program
- Helps training of volunteers
- Provides data to help institute change
- A method to support and factual for any changes which should be considered

The Role of Program Evaluation: Words/Ideas Cluster #7

Labeled and defined as enhance communication with multiple stakeholders (e.g., staff, public, program participants, other organizations, policy makers) through sharing of program results, marketing strategies, and reciprocal feedback opportunities with program participants:

- Demonstrate need for program to multiple constituents
- Helps communicate their benefits/impact to stakeholders
- Communicate results with funders, elected officials, and public
- Share with stakeholders about orgs impact and value (relevant, trusted)

- ID strengths and providing the ability to share that info with organization members and funders
- To have empirical data for funders
- Help in communication with public and policy advocacy
- For policy makers
- Briefing papers
- To help other organizations improve practice
- Report to referral sources to demonstrate value of collaboration
- Helps other agencies to know what works so more people can be helped
- For fellow providers
- Generate collaboration within organization and externally
- As sources of info for marketing
- Marketing materials and websites
- For potential clients
- To obtain realistic and honest feedback from the clients
- Customer satisfaction
- Give participants a voice
- Report to clients to engage them in program design
- Evidence of results for participants
- Client feedback: obtaining perspective from those using services.

The Role of Program Evaluation: Words/Ideas Cluster #8

Labeled and defined as cultivate funding collaborations such as providing data regarding the efficacy of program(s), providing information learned from evaluations, providing information regarding program improvements/development, satisfying funder requirements, and a means for garnering financial support:

- Report to funders and supporters to demonstrate impact of their contributions
- To share success with funders and show what the organization is doing to improve
- Convey to supporters program is successful
- Evidence for donors
- Proving to supporters that programs are relevant
- Fundable
- Helps make the case for support to donors
- What funders get for their contribution
- Donor should look for organizations continuously seeking improvement
- Generating resources if program is proven successful
- Provide information to funders
- Provide supporters with information
- To have empirical data for policy maker funders
- Help in talking with funders
- Meeting funders requirement
- For funders
- Satisfy funders
- Fund raise
- Adding to data to the case for support-assist in fund development efforts
- Secure more funding with independent evaluation or good objective internal evaluation
- Helps to get funding

- Enhances their resources and capacity to continue to provide services General additional resources
- Produce data that gives support for the need of the program
- Obtain new or additional resources
- Generating resources if the program is proven successful

Appendix M

Words/Ideas Clusters for Essential Capacity Elements

The following eight clusters of words and ideas came from the Delphi panelists responses from the first exploratory survey round. The responses were generated from three questions related to the capacity for program evaluation:

- Q6: What are the main challenges the nonprofit sector in Northeast Florida faces regarding program evaluation?
- Q7: What (if any) are the resources needed to develop capacity in order to sustain program evaluation efforts in the Northeast Florida nonprofit sector?
- Q8:How do funders and providers in the Northeast Florida nonprofit sector work together to develop program evaluation strategies (e.g., objectives, outcomes, evaluation purposes, uses of evaluation results, evaluation approaches, capacity issues, resource allocations)?

Q6 is a challenge question and responses to that were clustered with responses to the other questions that were similar.

Capacity Elements: Words/Ideas Cluster #1

Labeled and defined as sufficient time available for staff to plan, implement, analyze, and reflect on results from program evaluation:

- Lack of staff time to dedicate to the preparation and execution of the evaluation
- Time consuming task that takes significant resources to accomplish
- Limited time to conduct and to study and utilize evaluations
- Staff time available for evaluations
- The costs in terms of time
- Time

- Staff time
- The push to evaluate, evaluate, evaluate is great…if it's used well. If not, it can waste valuable time that nonprofits could be serving clients.
- Resources: time
- Commitment of staff time…to conduct program evaluations and to discuss the outcomes to improve organizational program performance
- Good resources…time…for internal evaluations to work right. Hard to make time for deep and thorough evaluation

Capacity Elements: Words/Ideas Cluster #2

Labeled and defines as sufficient financial resources specifically designated for program evaluation efforts:

- Financial resources (x2)
- Financial constraints that make it difficult to incorporated lessons learned from evaluation into organization practice.
- Funders expect it but won't fund it
- The cost of doing evaluations with control groups or comparison groups
- Limited budget to do evaluations
- The cost in term of budget
- Money to help facilitate evaluation programs. So many granters want quality evaluations built into their funded program but provide NO funds to do evaluations.
- Funding for staff to complete program evaluations
- Commitment of financial resources
- Good resources…financial…for internal evaluations to work right

- Money for good external evaluations is hard to come by
- Money
- Funding: program evaluation adds expense, and often donors don't build these expenses into their gifts/grants.
- Funding included in programming grants for evaluation.
- Funders, provide money to ensure evaluations can be completed on projects or programs
- the question represents lack of understanding about funding constraints

Capacity Elements: Words/Ideas Cluster #3

Labeled and defined as sufficient human resources such as skilled and designated evaluation staff or access to/partnerships with experts (professional evaluators, researcher community, program evaluation technical assistance consultants, etc…):

- Access to expertise
- Access to expertise-professional evaluators
- Technical assistance to identify the best techniques for a particular program
- Build relationship between nonprofit center and research community to consider what should be evaluated and why
- A community expert that will be available to nonprofits for free to help staff develop evaluations on a "shoestring budget"
- For those without a dedicated evaluation staff person, a consultant that could design evaluation tools for their organization.
- No or little relationship with research community
- Funders provide technical assistance
- Knowledge/expertise

- Good resources for internal evaluations to work right
- Dedicated staff to carry out evaluations
- Doing it well requires a level or expertise that only larger organization may have
- The cost in terms of staff
- Experienced and skilled staff & volunteers
- Staffing
- Resources: human capital
- If it's internal, having the expertise and resources to create, implement, and interpret the results of a good evaluation process.

Capacity Elements: Words/Ideas Cluster # 4

Labeled and defined as a positive culture (organizational and sector-wide) that advances program evaluation efforts such as prioritizing resources, willingness to accept feedback, openness to change, and a readiness to reflect and act on results:

- The understanding of the importance of program evaluation
- Prioritizing the time, money, and effort required to complete evaluations
- Willingness to accept feedback and make constructive changes, qualitatively and financially
- Being open to identifying best performers and what that means for others (i.e. using program evaluation to facilitate collaborations or even mergers)
- Lack of desire to test assumptions
- Organizational commitment to the importance of program evaluations
- An agreement/commitment among the nonprofit sector that program evaluations are necessary.

- Desire to continuously improve results
- Building internal evaluation capacity
- Demonstration that it makes a difference for funding and policy
- Sector wide understanding of the importance to tracking outcomes and not just outputs

Capacity Elements: Words/Ideas Cluster #5

Labeled and defined as realistic expectations from the philanthropic community regarding program evaluation efforts such as evaluation design, expected outcomes, reporting requirements, and their response to the results:

- Sometimes forced by donors to track data that doesn’t help the organization understand program effectiveness or that organization has great difficulty gathering from other organizations
- Some funders are more subjective than others. All program evaluations should be based on measurable outcomes.
- Creating ways of measuring without creating unrealistic expectations: you can’t expect 6 or even 12 months of intervention to undo years
- Funders who create pressure for positive evaluations rather than creating an environment where even unfavorable or underwhelming evaluations can be used for learning
- Greater flexibility from donors-especially [name of funder omitted for confidentiality]
- I think there is a challenge for funders to devise a way for all of the organizations that they serve to provide information in a uniform way that will allow the programs to be compared to one another
- A number of funders are interested in numbers served and hours of program provided

- Each funder uses different tools for program evaluation, although many of the questions are the same/similar
- Everyone has different reporting requirements. I imagine its very time consuming for providers.
- Education for the philanthropic community, ideally through the Donors Forum or other venues, on what they should be requesting of the nonprofits they fund and figuring out how to "right size" program evaluation requests

Capacity Elements: Words/Ideas Cluster #6

Labeled and defined as functional program evaluation designs/methods that take into account feasibility, appropriate measurement tools, access to data, infrastructure to accommodate data, and issues related to engaging program participants in the evaluative process:

- If it's external, making sure the evaluators really understand the program and can develop the appropriate tools to help the organization get the results it needs
- The emphasis on evaluation in the nonprofit and philanthropic sectors right now, combined with confusion about evaluation…leads both nonprofits and funders to try to evaluate all peoples and organizations in the same way.
- Maintaining objectivity in designing the program evaluation tool
- Not knowing who should do the evaluation
- Evaluations or evaluators who lose sight of "real world" practice
- A challenge for us is that we have several curricula within programs, all of which have varying evaluation components
- Not understanding what should be measured

- The changing external environment, which sometimes makes last year's evaluation no longer relevant to this year's issues
- Creating tools to use in their evaluations
- Correct evaluation tools
- Identify common performance measures
- Measuring outputs rather than outcomes
- Look for data access solutions together
- Access to data and analytic tools
- Access to data
- Limited access or availability of data (no baseline data, biases, etc…)
- Solid infrastructure to accommodate the data
- A major technological infrastructure that links all programs using single client ID would allow incredible insights. This would be a real shift in thinking as it relates to things like confidentiality.
- What will be done with the final evaluation and who to provide the information to
- Response rates from those being evaluated
- Survey, survey, survey fatigue for participants
- For small organizations, to evaluate our program compared to constituents who did not participate
- Long-term contact with participants

Capacity Elements: Words/Ideas Cluster #7

Labeled and defined as ongoing collaboration in and between the philanthropic and provider communities through negotiating appropriate evaluation strategies, sharing resources, providing results, and opportunities for discussion:

- Depends on the funder and depends on the provider (in response to a question on level of collaboration)
- In competitive grants, providers are asked to define program evaluation elements in great detail then have two-way conversation at site visit to increase understanding
- Two-way conversation with negotiated grants where funders and providers discuss most appropriate evaluation strategies and define together.
- Nonprofit center is the go to place for all things nonprofit, but we are not collaborating around evaluation
- Florida philanthropic network (noted as an example of collaboration)
- Donors forum of NEFL (noted as an example of collaboration)
- United Way of Northeast Florida does yearly session to discuss what they would like to see in the reporting for that year (noted as an example of collaboration)
- United Way has a significant impact on our evaluation processes and prescribes many of the outcomes we measure. They are keenly interested in the processes we use, the appropriateness of sample size… (noted as an example of collaboration)
- The best funders are foundations, who are more likely to allow our agency to determine our own evaluation methods
- We worked with cohorts around reflective practice
- Homeless coalition has tried to bring together groups to evaluate if the sector is really helping people out of homelessness, and how to target chronological homelessness (noted as an example of collaboration)

- Children's commission and United Way (also duPont) work to coordinate evaluation processes when applicable
- Agree and appoint a coordinator to collect and distribute data
- Agree on a common objective and the purpose of evaluation
- I see very little collaboration between providers in terms of data collection. In some instances there are issues of confidentiality that might prevent sharing
- Not very well (in response to a question on collaboration)
- Share resources and evaluation methods
- Share results with other nonprofits and funders
- Share results information with each other
- Informal talking with one another

Capacity Elements: Words/Ideas Cluster #8

Labeled and defined as ongoing training for providers and funders regarding program evaluation strategies/best practices and available resources:

- Training/knowledge in program evaluation techniques
- Classes/sessions through organization like the United Way or Nonprofit Center on how to start program evaluation within your organization
- A training on how to use the data that the organization is already collecting to answer program evaluation questions
- Additional training through the Nonprofit Center (potentially in partnership with JCCI) to train nonprofit executives, boards, and program managers
- Educate funder community about costs and benefits of evaluation

- Training for staff and volunteers. United Way offers support when they change a process, but I am not aware of other opportunities for organization to learn how to do evaluation effectively.
- Training for staff
- Training for all staff
- Training provided through the Nonprofit Center
- Funders refer towards resources to increase knowledge (funders refer nonprofits)
- Inconsistent understanding of what evaluation is, its purpose, and its best uses-and confusion about terms like measurements, metrics, assessment, and evaluation

References

Adler, M., & Ziglio, E. (Eds.). (1996). *Gazing into the oracle: The Delphi method and its application to social policy and public health*. London, England: Jessica Kingsley Publishers.

Ahmed, Z., Ahmed, Z., Nawaz, M. M., Dost, K. B., & Khan, M. A. (2011). Comparative significance of the four perspectives of balanced scorecard. *Interdisciplinary Journal of Contemporary Research in Business, 3*(1), 981-994.

Alaimo, S. P. (2008). Nonprofits and evaluation: Managing expectations from the leader's perspective. In J. Carman & K. Fredericks (Eds.), Nonprofits and evaluation. *New Directions for Program Evaluation, 2008*(119), 73-92. doi: 10.1002/ev.269

American Educational Research Association. (2006). Standards for reporting on empirical social science research in AERA publications. *Educational Researcher, 35*(6), 33-40.

Argyris, C. (1999). *On organizational learning* (2nd ed.).Oxford, UK: Blackwell Publishers.

Armstrong, J. S. (2001). Selecting forecasting methods. In J. S. Armstrong (Ed.), *Principles of forecasting: A handbook for researchers and practitioners* (pp. 365–386). Norwell, MA: Kluwer Academic Publishers.

Arnett, J. C. (1993). Evaluation approaches and methods defined. In Council on Foundations (Ed.), *Evaluation for foundations: Concepts, cases, guidelines, and resources* (pp. 267-287). San Francisco, CA: Jossey-Bass.

Arnold, M. E. (2006). Developing evaluation capacity in extension 4-H field faculty. *American Journal of Evaluation, 27,* 257-269. doi:10.1177/1098214006287989

Atkinson, D. D., Wilson, M., & Avula, D. (2005). A participatory approach to building capacity of treatment programs to engage in evaluation. *Evaluation and Program Planning, 28,* 329-334. doi:10.1016/j.evalprogplan.2005.04.012

Axelrod, N. R. (2005). Board leadership and development. In R. D. Herman (Ed.), *The Jossey-Bass handbook of nonprofit leadership & management* (2nd ed., pp. 131-152). San Francisco, CA: Jossey-Bass.

Baizerman, M., Compton, D. W., & Stockdill, S. H. (2002). New directions for ECB. *New Directions for Evaluation, 2002*(93), 109–120. doi:10.1002/ev.45

Baruch, Y., & Ramalho, N. (2006). Communalities and distinctions in the measurement of organizational performance and effectiveness across for-profit and nonprofit sectors. *Nonprofit and Voluntary Sector Quarterly, 35,* 39-65. doi:10.1177/0899764005282468

Behrens, T. R., & Kelly, T. (2008). Paying the piper: Foundation evaluation capacity calls the tune. *New Directions for Evaluation*, *2008*(119), 37-50. doi:10.1002/ev.267

Belief. (2013). In *Merriam-Webster's* online dictionary. Retrieved from http://www.merriam-webster.com/thesaurus/belief

Bozzo, S. L. (2002). Evaluation capacity building in the voluntary/nonprofit sector. *The Canadian Journal of Program Evaluation*, *17*(3), 75-92.

Bradshaw, P., Murray, V., & Wolpin, J. (1992). Do nonprofit boards make a difference? An exploration of the relationships among board structure, process and effectiveness. *Nonprofit and Voluntary Sector Quarterly*, *21*, 227-249. doi:10.1177/089976409202100304

Brinckerhoff, P. C. (2000). *Mission-based management: Leading your not-for-profit in the 21st century* (2nd ed.). New York, NY: John Wiley & Sons, Inc.

Brown, R., & Reed, C. S. (2002). An integral approach to evaluating outcome evaluation training. *The American Journal of Evaluation*, *23*, 1-17. doi:10.1016/S1098-2140(01)00161-8

Brown, W. A. (2005). Exploring the association between board and organizational performance in nonprofit organizations. *Nonprofit Management & Leadership*, *15*, 317-339. doi:10.1002/nml.71

Bryson, J. M. (2004). *Strategic planning for public and nonprofit organizations: A guide to strengthening and sustaining organizational achievement* (3rd ed.). San Francisco, CA: Jossey-Bass.

Bryson, J.M. (2010). Strategic planning and the strategy change cycle. In D. O. Renz (Ed.), *The Jossey-Bass handbook of nonprofit leadership and management* (3rd ed., pp. 230-261). San Francisco, CA: Jossey-Bass.

Bryson, J. M., & Patton, M. Q. (2010). Analyzing and engaging stakeholders. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer (Eds.), *Handbook of practical program evaluation* (3rd ed., pp. 30-54). San Francisco, CA: Jossey-Bass.

Candler, G., & Dumont, G. (2010). A non-profit accountability framework. *Canadian Public Administration*, *53*, 259-279. doi:10.1111/j.1754-7121.2010.00126.x

Carman, J. G. (2007). Evaluation practice among community-based organizations. *American Journal of Evaluation*, *28*, 60-75. doi:10.1177/1098214006296245

Carman, J. G. (2010). The accountability movement: What's wrong with this theory of change? *Nonprofit and Voluntary Sector Quarterly*, *39*, 256-274. doi:10.1177/0899764008330622

Carman, J. G. (2011). Understanding evaluation in nonprofit organizations. *Public Performance & Management Review*, *34*, 350-377. doi:10.2753/PMR1530-9576340302

Carman, J. G., & Fredericks, K. A. (2008). Nonprofits and evaluation: Empirical evidence from the field. *New Directions for Evaluation*, *2008*(119), 51-71. doi:10.1002/ev268

Carman, J. G., & Fredericks, K. A. (2009). Evaluation capacity and nonprofit organizations: Is the glass half-empty or half-full? *American Journal of Evaluation*, *31*, 84-104. doi:10.1177/1098214009352361

Carver, J. (1997). *Boards that make a difference: A new design for leadership in nonprofit and public organizations* (2nd ed.). San Francisco, CA: Jossey-Bass.

Carver, J., & Carver, M. M. (1997). *Reinventing your board: A step-by-step guide to implementing policy governance*. San Francisco, CA: Jossey-Bass.

Carver, J., & Carver, M. (2009). The policy governance model. *Board Leadership: Policy Governance in Action*, 101, 1-8. doi:10.1002/bl.101

Chait, R., Ryan, W. P., & Taylor, B. E. (2005). *Governance as leadership: Reframing the work of nonprofit boards*. Hoboken, NJ: John Wiley & Sons.

Charity Navigator. (n.d.) How do we plan to evaluate results reporting? Retrieved on 5/19/13 from http://www.charitynavigator.org/index.cfm?bay=content.view&cpid=1507

Chaytor, K., MacDonald, G., & Melvin, D. (2002). Preparing nonprofits for new accountability demands. *The Canadian Journal of Program Evaluation*, *17*(3), 93-112.

Chelimsky, E. (1997). The coming transformations in evaluation. In E. Chelimsky & W. R. Shadish (Eds.), *Evaluation for the 21st century: A handbook* (pp. 1-26). Thousand Oaks, CA: Sage.

Chelimsky, E., & Shadish, W. R. (Eds.). (1997). *Evaluation for the 21st century: A handbook*. Thousand Oaks, CA: Sage.

Chen, C. C., & Jones, K. (2009). Are employees buying the balanced scorecard? *Management Accounting Quarterly*, *11*(1), 36-45.

Clibbens, N., Walters, S., & Baird, W. (2012). Delphi research: Issues raised by a pilot study. *Nurse Researcher*, *19*(2), 37-44.

Coe, C. K. (2007). *Governmental and nonprofit financial management*. Vienna, VA: Management Concepts.

Compton, D. W., Glover-Kudon, R., Smith, I. E., & Avery, M. E. (2002). Ongoing capacity building in the American Cancer Society (ACS) 1995– 2001. *New Directions for Evaluation*, *2002*(93), 47-62. doi:10.1002/ev.41

Cook, S. D. N., & Yanow, D. (2005). Culture and organizational learning. In J. M. Shafritz, J. S. Ott, & Y. S. Jang (Eds.), *Classics of organization theory* (6th ed., pp. 368-382). Boston, MA: Thomson-Wadsworth (Original work published 1993).

Coryn, C. L., Schroter, D. C., & Hanssen, C. E. (2009). Adding a time-series design element to the success case method to improve methodological rigor: An application for nonprofit program evaluation. *American Journal of Evaluation*, *30*, 80-92. doi:10.1177/1098214008326557

Cousins, J. B., Elliott, C., Amo, C., Bourgeois, I., Chouinard, J., Goh, S. C., & Lahey, R. (2008). Organizational capacity to do and use evaluation: Results of a pan Canadian survey of evaluators. *The Canadian Journal of Program Evaluation*, *23*(3), 1-35.

Cutt, J., & Murray, V. (2000). *Accountability and effectiveness evaluation in non-profit organizations*. New York, NY: Routledge.

Crutchfield, L. R., & Grant, H. M. (2008). *Forces for good: The six practices of high-impact nonprofits*. San Francisco, CA: Jossey-Bass.

Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science, 9*(3), 458-467.

Davidson, E. J. (2005). *Evaluation methodology basics: The nuts and bolts of sound evaluation*. Thousand Oaks, CA: Sage.

Day, L. H. (2002). Delphi research in the corporate environment. In H. A. Linstone & M. Turoff (Eds.), *The Delphi method: Techniques and applications* (pp. 162-188). Retrieved from http://is.njit.edu/pubs/delphibook/delphibook.pdf

Donmoyer, R. (1990). Generalizability and the single-case study. In E. W. Eisner & A. Peshkin (Eds.), *Qualitative inquiry in education* (pp. 175-200). New York, NY: Teachers College Press.

Duignan, P. (2003). Mainstreaming evaluation or building evaluation capability? Three key elements. *New Directions for Evaluation*, *2003*(99), 7-21. doi:10.1002/ev.90

Dugan, M. A. (1996). Participatory and empowerment evaluation: Lessons learned in training and technical assistance. In D. Fetterman, S. J. Kaftarian, & A. Wandersman (Eds.), *Empowerment evaluation: Knowledge and tools for self-assessment & accountability* (pp. 277-303). Thousand Oaks, CA: Sage.

Ebrahim, A. (2010). The many faces of nonprofit accountability. In D. O. Renz (Ed.), *The Jossey-Bass handbook of nonprofit leadership and management* (3rd ed., pp. 101-121). San Francisco, CA: Jossey-Bass.

Eckerd, A., & Moulton, S. (2011). Heterogeneous roles and heterogeneous practices: Understanding the adoption and uses of nonprofit performance evaluations. *American Journal of Evaluation*, *32*, 98-117. doi:10.1177/1098214010381780

Eisner, E. W. (1998). *The enlightened eye: Qualitative inquiry and the enhancement of educational practice*. Upper Saddle River, NJ: Prentice-Hall, Inc.

Ekionea, J.-P. B., & Fillion, G. (2011). Knowledge management capabilities consensus: Evidence from a Delphi study. *Academy of Information and Management Sciences Journal*, *14*(1), 25-52.

Fayol, H. (2005). General principles of management. In J. M. Shafritz, J. S. Ott, & Y. S. Jang (Eds.), *Classics of organizational theory* (6th ed., pp. 48-60). Boston, MA: Thomson-Wadsworth (Original work published 1916).

Febey, K., & Coyne, M. (2007). Program evaluation: The board game: An interactive learning tool for evaluators. *American Journal of Evaluation*, *28*, 91-101. doi:10.1177/1098214006296567

Ferguson, S. D. (2000). *Researching the public opinion environment: Theories and methods*. Thousands Oak, CA: Sage.

Fetterman, D. M. (1996). Empowerment evaluation: An introduction to theory and practice. In David. M. Fetterman, S. J. Kaftarian, & A. Wandersman (Eds.), *Empowerment evaluation: Knowledge and tools for self-assessment & accountability* (pp. 3-46). Thousand Oaks, CA: Sage.

Flagg, J. & Rankin, A. (2011). *Public attitudes towards nonprofit organizations.* Unpublished report, Nonprofit Center of Northeast Florida, Jacksonville, Florida.

Fletcher, J. K., & Käufer, K. (2003). Shared leadership: Paradox and possibility. In C. L. Pearce & J. A. Conger (Eds.), *Shared leadership: Reframing the hows and whys of leadership* (pp. 21-47). Thousand Oaks, CA: Sage.

Forbes, D. P. (1998). Measuring the unmeasurable: Empirical studies of nonprofit organization effectiveness from 1977 to 1997. *Nonprofit and Voluntary Sector Quarterly*, *27*, 183-202. doi:10.1177/0899764098272005

Franklin, K. K., & Hart, J. K. (2006). Idea generation and exploration: Benefits and limitations of the policy Delphi research method. *Innovative Higher Education*, *31*, 237–246. doi:10.1007/s10755-006-9022-8

Friedman, M. (2005). *Trying hard is not good enough: How to produce measurable improvements for customers and communities*. Bloomington, IN: Trafford Publishing.

Funnel, S. C., & Rogers, P. J. (2011). *Purposeful program theory: Effective use of theories of change and logic models*. San Francisco, CA: Jossey-Bass.

Garcia-Iriarte, E., Suarez-Balcazar, Y., Taylor-Ritzler, T., & Luna, M. (2011). A catalyst-for-change approach to evaluation capacity building. *American Journal of Evaluation, 32*, 168-182. doi:10.1177/1098214010387114

Geist, M. R. (2010). Using the Delphi method to engage stakeholders: A comparison of two studies. *Evaluation and Program Planning*, *33*, 147-154. doi:10.1016/j.evalprogplan.2009.06.006

Gill, M., Flynn, R. J., & Reissing, E. (2005). The governance self-assessment checklist: An instrument for assessing board effectiveness. *Nonprofit Management and Leadership, 15*, 271-294. doi:10.1002/nml.69

Gronbjerg, K. A, Liu, H. K., & Pollak, T. H. (2009). Incorporated but not IRS-registered: Exploring the (dark) grey fringes of the nonprofit universe. *Nonprofit and Voluntary Sector Quarterly*, *39*, 925-945. doi:10.1177/0899764009342898

Gulick, L. (2005). Notes on the theory of organization. In J. M. Shafritz, J. S. Ott, & Y. S. Jang (Eds.), *Classics of organization theory* (6th ed., pp. 79-87). Boston, MA: Thomson-Wadsworth (Original work published 1937).

Gupta, U. G., & Clarke, R. E. (1996). Theory and applications of the Delphi technique : A bibliography (1975-1994). *Technological Forecasting and Social Change*, *53*, 185-211.

Hall, P. D. (2010). Historical perspectives on nonprofit organizations in the United States. In D. O. Renz (Ed.), *The Jossey-Bass handbook of nonprofit leadership and management* (3rd ed., pp. 3–41). San Francisco, CA: Jossey-Bass.

Handy, F., Seto, S., Wakaruk, A., Mersey, B., Mejia, A., & Copeland, L. (2010). The discerning consumer: Is nonprofit status a factor? *Nonprofit and Voluntary Sector Quarterly*, *39*, 866-883. doi:10.1177/0899764010362113

Hasson, F., & Keeney, S. (2011). Enhancing rigour in the Delphi technique research. *Technological Forecasting and Social Change*, *78*, 1695-1704. Elsevier Inc. doi:10.1016/j.techfore.2011.04.005

Hendricks, M., Plantz, M. C., & Pritchard, K. J. (2008). Measuring outcomes of United Way-funded programs: Expectations and reality. *New Directions for Evaluation*, *2008*(119), 13-35. doi:10.1002/ev266

Herman, R. D. (2010). Executive leadership. In D. O. Renz (Ed.), *The Jossey-Bass handbook of nonprofit leadership & management* (3rd ed., pp. 157-177). San Francisco, CA: Jossey-Bass.

Herman, R. D., & Heimovics, D. (2005). Executive leadership. In R. D. Herman (Ed.), *The Jossey-Bass handbook of nonprofit leadership and management* (2nd ed., pp. 153-170). San Francisco, CA: Jossey-Bass.

Herman, R. D., & Renz, D. O. (1999). Theses on nonprofit organizational effectiveness. *Nonprofit and Voluntary Sector Quarterly*, *28*, 107-126. doi:10.1177/0899764099282001

Hopkins, B. R., & Gross, V. C. (2010). The legal framework of the nonprofit sector in the United States. In D. O. Renz (Ed.), *The Jossey-Bass handbook of nonprofit leadership and management* (3rd ed., pp. 42-76). San Francisco, CA: Jossey-Bass.

Howe, K., & Eisenhart, M. (1990). Standards for qualitative (and quantitative) research: A prolegomenon. *Educational Researcher*, *19*(4), 2-9. doi:10.2307/1176381

Hung, H.-L., Altschuld, J. W., & Lee, Y.-F. (2008). Methodological and conceptual issues confronting a cross-country Delphi study of educational program evaluation. *Evaluation and Program Planning*, *31*, 191-198. doi:10.1016/j.evalprogplan.2008.02.005

Iecovich, E., & Bar-Mor, H. (2007). Relationships between chairpersons and CEOs in nonprofit organizations. *Administration in Social Work*, *31*(4), 21-41. doi:10.1300/J147v31n04_04 21

Internal Revenue Code, 26 U.S.C. § 501 (2010).

Jackson, D. K., & Holland, T. P. (1998). Measuring the effectiveness of nonprofit boards. *Nonprofit and Voluntary Sector Quarterly*, *27*, 159-182. doi:10.1177/0899764098272004

Johnson, B., & Christensen, L. (2008). *Educational research: Quantitative, qualitative, and mixed approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Johnson, R. M. (1993). Part two: Nine case studies of foundation evaluations. In Council on Foundations (Ed.), *Evaluation for foundations: Concepts, cases, guidelines, and resources* (pp. 45-265). San Francisco, CA: Jossey-Bass.

Kaplan, R. S. (2001). Strategic performance measurement and management in nonprofit organizations. *Nonprofit Management & Leadership*, *11*, 353-370. doi:10.1002/nml.11308

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard--measures that drive performance. *Harvard Business Review*, *70*(1), 71-79.

Kaplan, S. A., & Garrett, K. (2005). The use of logic models by community-based initiatives. *Evaluation and Program Planning*, *28*, 167-172. doi:10.1016/j.evalprogplan.2004.09.002

Karvonen, M., Ryynanen, V., & Kassi, T. (2009). The Delphi method as a tool for analyzing technology evolution: Case open source thin computing. *International Journal of Innovation and Technology Management*, *6*, 327-340. doi:10.1142/S0219877009001686

Kegeles, S. M., Rebchook, G. M., & Tebbetts, S. (2005). Challenges and facilitators to building program evaluation capacity among community-based organizations. *AIDS Education and Prevention*, *17*, 284-299. doi:10.1521/aeap.2005.17.4.284

Kehrer, B. H. (1993). Part I: Introduction and overview of evaluation. In Council on Foundations (Ed.), *Evaluation for foundations: Concepts, cases, guidelines, and resources* (pp. 1-43). San Francisco, CA: Jossey-Bass.

King, J. A. (2002). Building the evaluation capacity of a school district. *New Directions for Evaluation*, *2002*(93), 63-80. doi:10.1002/ev.42

Knauft, E. B., Berger, R. A., & Gray, S. T. (1991). *Profiles of excellence*. San Francisco, CA: Jossey-Bass.

Koppel, J. (2005). Pathologies of accountability: ICANN and the challenge of “Multiple accountabilities disorder.” *Public Administration Review*, 65(1), 94-108.

Kriemadis, T., Kotsovos, A., & Alexopoulos, P. (2009). Applying the balanced scorecard strategic evaluation method to a university athletic department. *Choregia*, *5*(2), 43-64. doi:http:dx.doi.org/10.4127/ch.2009.0040

Kvale, S. (1996). *Interviews: An introduction to qualitative research interviewing*. London, England: Sage.

Landeta, J. (2006). Current validity of the Delphi method in social sciences. *Technological Forecasting and Social Change*, *73*, 467-482. doi:10.1016/j.techfore.2005.09.002

Landeta, J., & Barrutia, J. (2011). People consultation to construct the future: A Delphi application. *International Journal of Forecasting*, *27*, 134-151. doi:10.1016/j.ijforecast.2010.04.001

Lee, Y.-F., Altschuld, J. W., & Hung, H.-L. (2008). Practices and challenges in educational program evaluation in the Asia-Pacific region: Results of a Delphi study. *Evaluation and Program Planning*, *31*, 368-375. doi:10.1016/j.evalprogplan.2008.08.003

Letts, C. W., Ryan, W. P., & Grossman, A. (1999). *High performance nonprofit organizations: Managing upstream for greater impact*. New York, NY: John Wiley & Sons, Inc.

Levi, D. (2007). *Group dynamics for teams* (2nd ed.). Thousand Oaks, CA: Sage.

Light, P. C. (2002). *Pathways to nonprofit excellence*. Washington, DC: Brookings Institution Press.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage. Retrieved from http://books.google.com/books?id=2oA9aWlNeooC&printsec=frontcover&dq=guba+and+lincoln+naturalistic+inquiry&source=bl&ots=0rnxW7RfBn&sig=_rcBtcLxl1TNHALQculzJ8ZixXg&hl=en&sa=X&ei=zZEmUJaMIoec8QT8roDYBw&ved=0CDkQ6AEwAA#v=onepage&q=guba and lincoln naturalistic inquiry&f=false

Linstone, H. A., & Turoff, M. (Eds.). (2002). *The Delphi method: Techniques and applications*. Retrieved from http://is.njit.edu/pubs/delphibook/delphibook.pdf

Littlepage, M. K., KBT & Associates. (2008). *Notes from the field: Strengthening a community's nonprofit sector*. Jacksonville, FL: Hartley Press.

Littlepage, M. K., & Oldakowski, R. (2006). *State of the sector: Duval County nonprofits* (2006 update). Retrieved from http://www.nonprofitctr.org/attachments/2006StateoftheNonprofitSector.pdf

Littlepage, M. K., & Oldakowski, R. (2012). *State of the sector: How Northeast Florida's nonprofits have changed and adapted in good times and bad* (Fall 2012). Retrieved from http://www.nonprofitctr.org/wp-content/uploads/2012/10/State-of-the-Sector.2012-UPDATE-FINAL-SM.pdf

Madaus, G. F., & Stufflebeam, D. L. (2000). Program evaluation: A historical overview. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (2nd ed., pp. 3–18). Norwell, MA: Kluwer Academic Publishers.

Mark, M. M., Henry, G. T., & Julnes, G. (2000). *Evaluation: An integrated framework for understanding, guiding, and improving public and nonprofit policies and programs*. San Francisco, CA: Jossey-Bass.

Marshall, C., & Rossman, G. B. (2006). *Designing qualitative research* (4th ed.). Thousand Oaks, CA: Sage.

Mayer, S. E. (1996). Building community capacity with evaluation activities that empower. In D. Fetterman, S. Kaftarian, & A. Wandersman (Eds.), *Empowerment evaluation: Knowledge and tools for self-assessment & accountability* (pp. 332-375). Thousand Oaks, CA.

McLaughlin, J. A., & Jordan, G. (1999). Logic models: A tool for telling your programs performance story. *Evaluation and Program Planning*, *22*, 65-72. doi:10.1016/S0149-7189(98)00042-1

Mendel, S. C. (2010). Are private government, the nonprofit sector, and civil society the same thing? *Nonprofit and Voluntary Sector Quarterly*, *39*, 717–733. doi:10.1177/0899764009337897

Miller, J. L. (2002). The board as a monitor of organizational activity: The applicability of agency theory to nonprofit boards. *Nonprofit Management and Leadership*, *12*, 429-450. doi:10.1002/nml.12407

Miller, R. L., & Campbell, R. (2006). Taking stock of empowerment evaluation: An empirical review. *American Journal of Evaluation*, *27*, 296-319. doi:10.1177/1098214006291015

Miller, T. I., Kobayashi, M. M., & Noble, P. M. (2006). Insourcing, not capacity building, a better model for sustained program evaluation. *American Journal of Evaluation*, *27*, 83-94. doi:10.1177/1098214005283185

Milstein, B., Chapel, T. J., Wetterhall, S. F., & Cotton, D. A. (2002). Building capacity for program evaluation at the Centers for Disease Control and Prevention. *New Directions for Evaluation*, *2002*(93), 27-46. doi:10.1002/ev.40

Murray, V. (2005). Evaluating the effectiveness of nonprofit organizations. In R. D. & A. Herman (Ed.), *The Jossey-Bass handbook of nonprofit leadership and management* (2nd ed., pp. 345-370). San Francisco, CA: Jossey-Bass.

Murray, V. (2010). Evaluating the effectiveness of nonprofit organizations. In D. O. Renz (Ed.), *The Jossey-Bass handbook of nonprofit leadership & management* (3rd ed., pp. 431-458). San Francisco, CA: Jossey-Bass.

Naccarella, L., Pirkis, J., Kohn, F., Morley, B., Burgess, P., & Blashki, G. (2007). Building evaluation capacity: Definitional and practical implications from an Australian case study. *Evaluation and Program Planning*, *30*, 231-236. doi:10.1016/j.evalprogplan.2007.05.001

National Center for Charitable Statistics. (n.d.). *Quick facts about nonprofits*. Retrieved from http://nccs.urban.org/statistics/quickfacts.cfm

Newcomer, K. E. (1997). Using performance measurement to improve programs. *New Directions for Evaluation*, *1997*(75), 5-14. doi:10.1002/ev.1076

Newcomer, K. E. (2001). Tracking and probing program performance: Fruitful path or blind alley for evaluation professionals? *American Journal of Evaluation*, *22*, 337-341. doi:10.1177/109821400102200308

Newcomer, K. E., Hatry, H. P., & Wholey, J. S. (2010). Planning and designing useful evaluations. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer (Eds.), *Handbook of practical program evaluation* (3rd ed., pp. 5-29). San Francisco, CA: Jossey-Bass.

Nobbie, P. D., & Brudney, J. L. (2003). Testing the implementation, board performance, and organizational effectiveness of the policy governance model in nonprofit boards of directors. *Nonprofit and Voluntary Sector Quarterly*, *32*, 571-595. doi:10.1177/0899764003257460

Nonprofit Center of Northeast Florida. (2013a). *Connecting nonprofits to information and resources.* Retrieved from http://www.nonprofitctr.org/

Nonprofit Center of Northeast Florida. (2013b). *Moving forward together*. Retrieved on 5/1/13 from http://www.nonprofitctr.org/convene/moving-forward-together/

Northeast Florida Regional Council (n.d.). *About us*. Retrieved on 4/11/13 from http://www.nefrc.org/about.htm

Ostrower, F. (2007). *Nonprofit governance in the United States: Findings on performance and accountability from the first national representative study* (Research Report No. 411479). Retrieved from Urban Institute website: http://www.urban.org/uploadedPDF/411479_Nonprofit_Governance.pdf

Ostrower, F. (2008). *A better way to deal with the leadership crisis* (Report No. 901174). Retrieved from the Urban Institute website: http://www.urban.org/url.cfm?ID=901174.

Paton, R. (2003). *Managing and measuring social enterprises*. Thousand Oaks, CA: Sage.

Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text* (3rd ed.). Thousand Oaks, CA: Sage.

Patton, M. Q. (1999). Organizational development and evaluation. *The Canadian Journal of Program Evaluation*, *14*(Special Issue), 93-113.

Patton, M. Q. (2002a). *Qualitative research & evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Patton, M. Q. (2002b). *Utilization-focused evaluation (U-FE) checklist*. Retrieved from http://www.wmich.edu/evalctr/archive_checklists/ufe.pdf

Patton, M. Q. (2007). Process use as a usefulism. *New Directions for Evaluation*, *2007*(116), 99-112. doi:10.1002/ev246

Patton, M. Q. (2010). Utilization-focused evaluation for social services and social work. *Social Work Review / Revista de Asistenta Sociala*, *2010*(4), 11-27.

Patton, M. Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. New York, NY: The Guilford Press.

Pearce, C. L., & Conger, J.A. (Eds.). (2003). *Shared leadership: Reframing the hows and why of leadership.* Thousand Oaks, CA: Sage.

Perception. (2013). In *Oxford's* online dictionary. Retrieved from http://oxforddictionaries.com/definition/english/perception

Peshkin, A. (1988). In search of subjectivity-One's own. *Educational Researcher, 17*(7), 17-21.

Plantz, M. C., Greenway, M. T., & Hendricks, M. (1997). Outcome measurement: Showing results in the nonprofit sector. *New Directions for Evaluation*, *1997*(75), 15-30. doi:10.1002/ev.1077

Poole, D. L., Nelson, J., Carnahan, S., Chepenik, N. G., & Tubiak, C. (2000). Evaluating performance measurement systems in nonprofit agencies: The Program Accountability Quality Scale. *American Journal of Evaluation*, *21*, 15-26. doi:10.1016/S1098-2140(00)00049-7

Posavac, E. J., & Carey, R. G. (2003). *Program evaluation: Methods and case studies* (6th ed.). Upper Saddle River, NJ: Pearson Education.

Preskill, H., & Torres, R. T. (1999). *Evaluative inquiry for learning in organizations*. Thousand Oaks, CA: Sage

Quinn, R. E., & Rohrbaugh, J. (1981). A competing values approach to organizational effectiveness. *Public Productivity Review*, *5*, 122-140.

Renz, D. O. (2010). Leadership, governance, and the work of the board. In D. O. Renz (Ed.), *The Jossey-Bass handbook of nonprofit leadership & management* (3rd ed., pp. 125-156). San Francisco, CA: Jossey-Bass.

Roeger, K. L., Blackwood, A., & Pettijohn, S. L. (2011). *The nonprofit sector in brief: Public charities, giving and volunteering, 2011* (Research Report No. 412434). Retrieved from Urban Institute website: http://www.urban.org/url.cfm?ID=412434

Schein, E. H. (2005). Defining organizational culture. In J. M. Shafritz, J. S. Ott, & Y. S. Jang (Eds.), *Classics of organization theory* (6th ed., pp. 360-367). Boston, MA: Thomson-Wadsworth (Original work published 1993).

Scriven, M. (1981). *Evaluation thesaurus* (3rd ed.). Inverness, CA: Edgepress.

Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization*. New York, NY: Currency Doubleday.

Sherlock, J. J., & Nathan, M. L. (2007). Nonprofit association CEOs: How their context shapes what, how, and why they learn. *Nonprofit Management & Leadership*, *18*, 19-39. doi:10.1002/nml169

Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The Delphi method for graduate research. *Journal of Information Technology Education, 6*, 1–21.

Sowa, J. E., Selden, S. C., & Sandfort, J. R. (2004). No longer unmeasurable? A multidimensional integrated model of nonprofit organizational effectiveness. *Nonprofit and Voluntary Sector Quarterly*, *33*, 711-728. doi:10.1177/0899764004269146

Stevenson, J., Florin, P., Mills, D. S., & Andrade, M. (2002). Building evaluation capacity in human service organizations: A case study. *Evaluation and Program Planning*, *25*, 233-243. doi:10.1016/S0149-7189(02)00018-6

Stockdill, S. H., Baizerman, M., & Compton, D. W. (2002). Toward a definition of the ECB process: A conversation with the ECB literature. *New Directions for Evaluation*, *2002*(93), 7-25. doi:10.1002/ev.39

Stufflebeam, D. L. (2000). The CIPP model for evaluation. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (2nd ed., pp. 279–317). Norwell, MA: Kluwer Academic Publishers.

Stufflebeam, D. L. (2001). Evaluation models [Monograph]. *New Directions for Evaluation, 2001*(89), 7-98. doi:10.1002/ev.3

Stufflebeam, D. L. (2007, March 17). *CIPP evaluation model checklist* (2nd ed.). Retrieved from http://www.wmich.edu/evalctr/archive_checklists/cippchecklist_mar07.pdf

Stufflebeam, D. L., & Wingate, L. A. (2005). A self-assessment for use in evaluation training. *American Journal of Evaluation*, *26*, 544-561. doi:10.1177/1098214005279730

Tang, H., Cowling, D. W., Koumjian, K., Roesler, A., Lloyd, J., & Rogers, T. (2002). Building local program evaluation capacity toward a comprehensive evaluation. *New Directions for Evaluation*, *2002*(95), 39-56. doi:10.1002/ev.57

Tapio, P., Paloniemi, R., Varho, V., & Vinnari, M. (2011). The unholy marriage? Integrating qualitative and quantitative information in Delphi processes. *Technological Forecasting and Social Change*, *78*, 1616-1628. doi:10.1016/j.techfore.2011.03.016

Taylor-Powell, E., & Henert, E. (2008). *Developing a logic model : Teaching and training guide*. Madison, WI: University of Wisconsin-Extension. Retrieved from http://www.uwex.edu/ces/pdande/evaluation/pdf/lmguidecomplete.pdf

Thomas, J. C. (2010). Outcome assessment and program evaluation. In D. O. Renz (Ed.), *The Jossey-Bass handbook of nonprofit leadership & management* (3rd ed., pp. 401-430). San Francisco, CA: Jossey-Bass.

Turoff, M., & Hiltz, S. R. (1996). Computer-based Delphi processes. In M. Adler & E. Ziglio (Eds.), *Gazing into the oracle: The Delphi method and its application to social policy and public health* (pp. 56-85). London, England: Jessica Kingsley Publishers.

Urban, J. B., & Trochim, W. (2009). The role of evaluation in research--practice integration working toward the "Golden Spike." *American Journal of Evaluation, 30,* 538-553. doi:10.1177/1098214009348327

Urban Institute (n.d.). *Nonprofits*. Retrieved from http://www.urban.org/nonprofits/index.cfm

U.S. Department of the Treasury, Internal Revenue Service. (2011). *Tax-exempt status for your organization* (IRS Publication No. 557). Retrieved from http://www.irs.gov/pub/irs-pdf/p557.pdf

U.S. Department of the Treasury, Internal Revenue Service. (2012). *Life cycle of an exempt organization.* Retrieved from http://www.irs.gov/charities/article/0,,id=169727,00.html

W. K. Kellogg Foundation. (2004a). *Evaluation handbook: Philosophy and expectations*. Retrieved from http://www.wkkf.org/knowledge-center/resources/2010/W-K-Kellogg-Foundation-Evaluation-Handbook.aspx

W. K. Kellogg Foundation. (2004b). *Logic model development guide*. Retrieved from http://www.wkkf.org/knowledge-center/resources/2006/02/WK-Kellogg-Foundation-Logic-Model-Development-Guide.aspx

Woodwell, W. H. (2005). *Evaluation as a pathway to learning: Current topics in evaluation for grantmakers*. Washington, D.C. Retrieved from http://www.geofunders.org/uploads/documents/live/GEOEvalRndTableReport05_FINAL.pdf

Worth, M. J. (2009). *Nonprofit management: Principles and practice*. Thousand Oaks, CA: Sage Publications, Inc.

Wu, J. Y. (2008). A general behavior model and new definitions of organizational cultures. *The Journal of Socio-Economics*, *37*, 2535-2545. doi:10.1016/j.socec.2007.10.002

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage

Ziglio, E. (1996). The Delphi method and its contribution to decision-making. In M. Adler & E. Ziglio (Eds.), *Gazing into the oracle: The Delphi method and its application to social policy and public health* (pp. 3-33). London, England: Jessica Kingsley Publishers.

Vita

Gail A. Patin, LCSW

SUMMARY OF QUALIFICATIONS

- Superb administrator with more than 20 years experience in supervision, program development, strategic planning, human resources and fiscal management.
- Dynamic leader, team builder, and problem solver through optimizing resources, prioritizing staff development, implementing positive delegation strategies, initiating joint decision making and consistently motivating others toward success.
- Proficient in grant development, management, and implementation. Experienced with federal, state and private foundation grant writing.
- Adept at public speaking, public relations and experienced with all media venues.
- Proven leader in coalition building between communities, service providers, military, law enforcement and government agencies through participation on strong coordinated community response initiatives. Established the Northeast Florida Task Force Against Domestic Violence. Key participant in other multi-disciplinary pilot projects, protocols, and committees.
- Launched and coordinate ground-breaking and nationally recognized projects such as the Community Partnership for the Protection of Children and the Intimate Violence Enhanced Services Team.
- Successful in cultivating relationships with individual donors, foundations, and government funders resulting in additional revenue streams.

EDUCATION

Doctorate of Education in Educational Leadership, University of North Florida

Master of Social Work, University of South Carolina

Bachelor of Science, Charleston Southern University

PROFESSIONAL EXPERIENCE

Hubbard House, Inc., Jacksonville, FL Chief Operating Officer	April 2006-Present
Harbor House of Central Florida, Orlando, FL Chief Executive Officer	March 2004 – April 2006
Hubbard House, Inc., Jacksonville, FL Chief Operating Officer	June 1995-March 2004
Children’s Home Society of Florida, Jacksonville, FL Senior Social Worker	October 1991- June 1995