

NONPROFIT LEADER PERCEPTIONS OF EFFECTIVE ORGANIZATIONAL  
PERFORMANCE MEASUREMENT: A Q METHODOLOGY STUDY

by

Diana Frayne

Copyright 2014

A Dissertation Presented in Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Management

University of Phoenix

UMI Number: 3648297

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3648297

Published by ProQuest LLC (2014). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code



ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 - 1346

The Dissertation Committee for Diana Frayne certifies approval of the following  
dissertation:


NONPROFIT LEADER PERCEPTIONS OF EFFECTIVE ORGANIZATIONAL  
PERFORMANCE MEASUREMENT: A Q METHODOLOGY STUDY

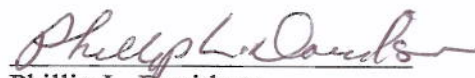
Committee:

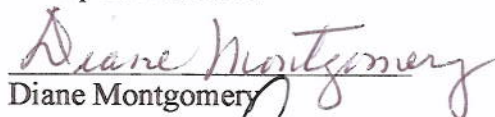
James Goes, PhD, Chair

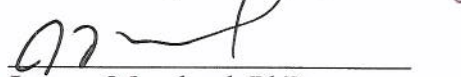
Phillip L. Davidson, PhD, Committee Member

Diane Montgomery, PhD, Committee Member

  
James Goes

  
Phillip L. Davidson

  
Diane Montgomery

  
Jeremy Moreland, PhD  
Dean, School of Advanced Studies  
University of Phoenix

Date Approved: June 27, 2014

## ABSTRACT

There is ample discourse regarding the need for changing nonprofit performance measurement, but there is little consensus within the field on how best to evaluate while maximizing performance, outcomes, and mission achievement. This Q methodology study documented the perceptions of 22 nonprofit leaders in the United States about effective performance measurement and the characteristics necessary to create an effective model to measure nonprofit performance. The study involved analyzing the nonprofit leaders' responses to create three distinct views on effective organizational performance measurement called (a) Road Map, (b) Management Tool, and (c) Weakest Link. Despite differences in the viewpoints, three themes emerged as a starting point to inform the shift in measuring nonprofit effectiveness: (a) the need for larger performance management systems, (b) eliminating the unfunded mandate for performance measurement, and (c) the desire for organization-specific mission-based outcome measurement. Insights from the nonprofit leaders revealed the characteristics of a new system for generating meaningful nonprofit performance data. The implementation of these characteristics could strengthen performance management, promote organizational learning, and inspire collaborative partnerships with funders and beneficiaries. Nonprofit leaders must create a culture of performance management that facilitates performance measurement and performance improvements if they are to advance the mission of the organizations they lead.

## DEDICATION

For the many nonprofit leaders who dedicate their time and energy to make the world a better place, one community at a time.

*Never doubt that a small group of thoughtful, committed citizens can change the world; indeed, it's the only thing that ever has.*

- Margaret Mead

## ACKNOWLEDGMENTS

This has been a long journey and it took many people to help forge my way to the completion of my research. First, I would like to acknowledge the love and support provided by my family and friends. I would like to thank my husband Jason Frayne for being my rock. Without his love, support, and encouragement of my insatiable curiosity and passion for learning, I would have never become the woman I am today. I am grateful that my mother, Karen Reed-Buell, taught me from a very young age the gift of giving back, even when the only thing you have to give is your time and effort. Only later did I realize that time and effort are the greatest gifts a person can truly give.

I would also like to acknowledge the gentle strength that my grandmother Rueda Reed exemplified and I try now to emulate, a characteristic that was tried and reinforced throughout this process. I am indebted to my sister Gloria Guy, as she was one of the few in my life who was able to truly empathize my experience; I hope to provide the same empathy during her own dissertation-writing journey. I am appreciative of my sister Teresa Guy, because she always helped to ground me and reminded me to laugh my way through each challenge. I would like to acknowledge my uncle Kym Reed for teaching me a strong work ethic, as it was the foundation I needed to reach the finish line. I would like to thank my aunt Leslie Reed for giving me the vocabulary for a great and flavorful life and for her indelible compassion. I am full of the aloha spirit thanks to my uncle Kris and aunt Kim Reed. I would be remiss not to thank my friends Jennifer NW Martin, Leilani Sampayan, Katie Perrin, and my many other friends for listening, even when I did not make sense, and for your encouragement and distractions along the way.

Second, I would like to acknowledge the guidance and support provided by my committee. Thank you to Dr. James Goes, my mentor and committee chair, for sticking with me during two moves, three new jobs, and the many transitions that stretched this process out longer than it should have taken. You challenged me to refine my ideas into a coherent and meaningful study. Thank you to Dr. Diane Montgomery for serving on my committee. Without your expertise, I would have been scrambling to understand Q methodology. And finally, thank you to Dr. Phil Davidson for serving on my committee and inspiring many insights in my writing. My committee's time, knowledge, guidance, and support were critical to my success.

Last, I must acknowledge my many mentors who helped guide me through this process. I am appreciative of my friend and colleague Carol Sommers for challenging me to start working on my doctoral degree. I may have had other words to say along the way, but I am thankful you started me down this road. I am grateful to my original mentor, Dr. Joyce Kaufman, as she helped me to craft an idea and passion into a career of giving. Thank you for believing in me even when I did not always believe in myself and for pushing me to always think critically and analytically. To my peer mentor Dr. William Dean, thank you for introducing me to Q methodology, for charging ahead and showing me the way, and for encouraging me to keep fighting to the finish.

## TABLE OF CONTENTS

Contents	Page
LIST OF TABLES .....	xiii
LIST OF FIGURES .....	xiv
Chapter 1 Introduction .....	1
Background .....	4
Issues in nonprofit organizational performance measurement .....	5
Widespread adoption of financial efficiency standards .....	5
Efforts to measure nonprofit organizational performance .....	8
Problem Statement .....	10
Purpose Statement .....	11
Significance .....	13
Significance of the study to the field .....	13
Significance of the study to leadership .....	16
Theoretical Framework .....	17
Systems theory .....	17
Resource dependence and institutional theories .....	18
Social constructionist theory .....	20
Research Questions .....	22
Nature of the Study .....	23
Definition of Terms .....	28
Assumptions .....	30
Scope .....	32



Limitations .....	33
Delimitations.....	35
Summary.....	36
Chapter 2 Literature Review.....	38
Documentation.....	40
Historical Overview .....	41
Theories of Nonprofit Performance .....	45
Systems theory .....	45
Resource dependence and institutional theories .....	47
Social constructionist theory.....	50
A change in expectations .....	54
Current Performance Measurement and Evaluation.....	55
Accountability and efficiency (external evaluation).....	57
Multidimensional evaluation .....	63
Balanced scorecard .....	65
Multidimensional, integrated model for nonprofit organizational effectiveness (MIMNOE) .....	65
Logic models.....	66
Social return on investment. ....	67
Outcome measurement (internal evaluation) .....	67
Empirical research .....	69
Limitations .....	72
Capacity building efforts.....	73

Mixed evaluation (insourcing) .....	74
High-impact and social value measurement. ....	75
Lack of consensus and disconnected momentum .....	76
What Is the Purpose of Performance Evaluation? .....	77
Creating Effective Performance and Performance Measurement.....	78
Internal characteristics .....	78
Systems thinking.....	79
Feedback mechanisms .....	79
Adaptive capacity.....	80
Appreciative inquiry .....	80
External characteristics .....	81
The leader’s importance to performance measurement .....	81
Perceptions of Organizational Effectiveness and Performance Measurement	
.....	82
Current related research .....	83
Performance in relation to leader characteristics .....	83
NPO leader perceptions of funding.....	84
Perceptions of program evaluation capacity .....	85
Leader perceptions captured while studying evaluation practices	
.....	87
Performance measurement characteristics .....	88
Performance measurement practices.....	89
An argument for leader perceptions of performance measurement	90

Review of Methodologies .....	91
Quantitative research methods .....	92
Qualitative research methods .....	94
Mixed-method research .....	95
Q methodology.....	96
Conclusion .....	96
Summary .....	98
Chapter 3 Method .....	100
Research Method .....	101
Appropriateness of Design.....	103
Q methodology.....	108
Study Population and Sample .....	113
Informed Consent.....	116
Confidentiality .....	117
Geographic Location.....	118
Instrumentation .....	118
Data Collection .....	119
Data Analysis .....	121
Summary.....	123
Chapter 4 Results .....	125
Overview of Data Collection .....	126
Description of Participants.....	127
Data Analysis .....	128

Interpretation of Factors.....	131
Research Question 1 .....	132
Factor 1: Road Map .....	133
Factor 2: Management Tool.....	139
Factor 3: Weakest link .....	144
Research Question 2 .....	149
Summary.....	153
Chapter 5 Conclusion and Recommendations .....	157
Summary of the Study .....	160
Discussion of Results.....	164
Performance measurement for Road Map Leaders.....	166
Performance measurement for Management Tool leaders. ....	167
Performance measurement for Weakest Link leaders. ....	169
Commonalities and implementation considerations. ....	170
Significance of the Study for leadership.....	175
Recommendations to Nonprofit Leaders .....	178
Limitations of the Study.....	182
Directions for Future Research .....	184
Summary of the Study .....	185
Concluding Remarks.....	188
References.....	194
Appendix A Participation Request Letter .....	225
Appendix B Informed Consent .....	227

Appendix C Q sample Statements .....	229
Appendix D Q sorting Instructions .....	234
Appendix E Pre- and Postexercise Questionnaire .....	236
Appendix F Participant Demographic Information .....	240
Appendix G Full Tables of Most Important/Most Unimportant Statements and Scores by Factor.....	246

LIST OF TABLES

Table 1 *Summary of Literature Searched by Category (1990-2012)* ..... 41

Table 2 *Factor Matrix With Defining Sorts* ..... 130

Table 3 *Correlations Between Pairs of Factors* ..... 131

Table 4 *Most Important Consensus Statements With Array Positions* ..... 150

Table 5 *Most Unimportant Consensus Statements With Array Positions*..... 140

Table 6 *Most Important and Most Unimportant Statements and Scores for Road Map  
Condensed*..... 144

Table 7 *Most Important & Most Unimportant Consensus Statements With Array  
Positions* ..... 150

LIST OF FIGURES

*Figure 1.* Quasi-normal Q sort matrix. .... 121

## Chapter 1

### Introduction

The nonprofit sector is growing rapidly and plays a prominent role in influencing social change (S. R. Smith, 2010). Between 1993 and 2008, the number of nonprofit organizations doubled to 1.9 million tax-exempt organizations and the number continues to rise (Independent Sector, 2012a). The missions of many nonprofit organizations are similar, creating broad overlap in service delivery and increased competition for funding, public attention, and volunteer support.

Funding levels in nonprofit organizations are decreasing, as more organizations are vying for declining philanthropic support. Projections for 2009 donations indicated a 4.8 to 5.7% decrease from 2008 donation levels, and projections for 2010 indicated an additional decline of 3.9% (Havens & Schervish, 2010). The benefits of donations are potentially negated because of the threat to organizational effectiveness from decreases in resource levels (Froelich, 1999). The forecast of charitable donations could become more competitive with potential tax changes looming. In May 2011, the Congressional Budget Office released a report with 11 recommendations for changes to the current charitable giving tax treatment; seven of the tax credit changes, if implemented, would decrease the amount of charitable giving available to the nonprofit sector (Congressional Budget Office, 2011; Independent Sector, 2011).

Government, donor, and public demand for the evaluation of nonprofit organizations have resulted in a challenge to create a comparative measure for strength and viability across the wide spectrum of differences represented in the nonprofit sector. For-profit organizations are generally evaluated and compared through financial



measures. This method of analysis has been applied to nonprofit organizations, but is not a true indicator of nonprofit effectiveness or potential. In the nonprofit setting, finances are the enabler of the mission, not the goal (Lindenberg, 2001). Leaders of for-profit organizations “need non-financial measures to help them find the *means* to achieve the *end* of remaining profitable. Non-profit managers . . . need non-financial measures to tell them whether they have used their financial resources as effective means for creating publically valuable results” (Moore, 2003, p. 8). Financial measures indicate how successful the organization has performed in the past and are not necessarily an indicator of future success (Kaplan, 2001).

Although donors and third-party organizations, such as private foundations and charity watchdogs, have evaluated nonprofit organizations on financial measures such as overhead and fundraising efficiency, there is a shift in dialogue among these groups, with nonprofit leaders and academics calling for new performance measures. The nonprofit sector might benefit from the development of a method to assess performance, effectiveness, and social impact, which has been a problematic endeavor as these factors are largely intangible and many believe are difficult, if not impossible, to compare from one organization to the next (Herman & Renz, 1999).

Leaders in the nonprofit sector are divided on how to evaluate nonprofit organizational performance most effectively and how to define organizational effectiveness (Lecy, Schmitz, & Swedlund, n.d.). Nonprofit organizations are subject to many stakeholders who often have conflicting expectations (Moore, 2003). These stakeholders include upstream customers, the donors who make nonprofit organizational work possible, and downstream customers, the clients and recipients of the social benefit

produced by nonprofit organizations. NPO evaluation (a) occurs more often to appease funders; (b) is defined and framed differently depending on the source, which creates an inconsistent definition of organizational effectiveness; and (c) is rarely used to change behavior or inspire organizational learning (Ebrahim, 2005). The greatest challenge is finding a balance among upward accountability to donors, meeting the needs of employees, and making positive change in organizational practices (Ebrahim, 2005).

Many theoretical and conceptual frameworks have been presented for measuring performance, yet few have gained significant backing or widespread adoption because few have yet to be empirically tested or statistically proven (Lacy, et. al, n.d.). Many models, systems, and subsystems examine different dimensions and subsections of performance. Each of these methods has inherent strengths, yet all fail to take a holistic view of nonprofit organizational performance, and few have been able to address performance across the nonprofit sector's diverse organizational demographics and missions.

With the rich and diverse field of opinions and methods, it is not surprising that empirical findings demonstrate the difficulty nonprofit organizations have in defining the purpose of organizational performance measurement in addition to designing and implementing effective performance measurement systems. The current Q methodology study examined nonprofit leader perceptions regarding effective nonprofit organizational performance measurement. The focus was on the subjective perceptions held by nonprofit leaders about which elements and characteristics are most important to creating an effective performance measurement model that may supplement or replace the current financial efficiency standard.

Chapter 1 contains an introduction to the problem and background of nonprofit organizational evaluation, in addition to the purpose of the research. The significance, theoretical framework, assumptions, scope, and limitations of the study are also reviewed. Through a review of the literature, Chapter 2 includes a foundation of background knowledge required to provide an explanation of how the current study fits into the historical, theoretical, and current body of research. Chapter 3 includes the plan and methodology for the current research, in addition to a discussion about the appropriateness of the design and research population. Chapter 4 includes detailed analysis of the Q sorts, descriptive demographic data, and clarification gained from post-sort interviews in relation to the research questions and the resulting factors. Last, Chapter 5 includes conclusions based upon the summary of results, limitations of the study, recommendations, significance of the research to leadership, and future implications concerning theory, practice, and research.

## **Background**

More than 1.9 million tax-exempt organizations exist in the United States, of which 1.4 million have 501(c)3 nonprofit organizational status with the Internal Revenue Service (IRS; Independent Sector, 2012a). Information available from 2009 indicated the nonprofit sector employed approximately 10% of the U.S. workforce, or 13.5 million people, and accounted for 5.5% of the gross domestic product (Independent Sector, 2012b). Americans donated \$295 billion and 12.9 billion hours, valued at \$215.6 billion, to charity in 2006 (Wing, Pollak, & Blackwood, 2008). Attention toward the sector as a valuable service provider has increased from the Obama Administration's Social Innovation Fund in addition to the passage of the Edward M. Kennedy Serve America

Act of 2009 (S. R. Smith, 2010). As one of the fastest growing sectors, growth in the nonprofit sector has caused increased scrutiny and demand for transparency from donors and the government (Ingram, 2009).

**Issues in nonprofit organizational performance measurement.** Limited resources and a growing demand for performance evaluation challenge many nonprofit organizations. Few have the personnel, the financial resources, or the training and tools necessary to conduct internal evaluations (Carman & Fredericks, 2010). Without proper capacity or resources, performance evaluation becomes tertiary to operations and fundraising, which is compounded by external attention focused on efficiency rather than performance outcomes (Ebrahim, 2005).

The evaluation measures most widely accepted and applied are financial indicators such as fundraising efficiency and overhead percentages (Wing, 2009). The American public has been trained to ask one simplified question of the nonprofit sector: How much of my donation goes to the cause (Pallotta, 2008)? This question relies on two ratios, fundraising efficiency and program spending, which are easily accessed and analyzed through IRS Form 990s (Pallotta, 2008; Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004a).

**Widespread adoption of financial efficiency standards.** Third-party evaluators, such as Charity Navigator, Better Business Bureau Wise Giving Alliance, and GuideStar, use various ranking systems, each grounded in financial performance derived from IRS Form 990s. Donors, charity watchdogs, and the media have vilified nonprofit organizations that exceed the acceptable overhead percentage set by third-party evaluators (Pallotta, 2008). Arbitrarily set between 65 and 75%, the efficiency standards

are not backed by research to indicate the numbers are true indicators of efficiency (Pallotta, 2008; Sloan, 2009; Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004a).

Nonprofit organizations have by necessity accepted the standards because their existence relies on the flow of donations, despite the fact that these artificial efficiency standards lower organizational capacity and the ability to perform the mission (Pallotta, 2008; Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004a). In an attempt to meet the efficiency standards, leaders of nonprofit organizations have used creative accounting methods for defining or hiding overhead, further skewing the public's understanding of the cost to operate and to serve the mission (U.S. General Accounting Office, 2002). Such creative accounting results from the IRS allowing charities ample discretion in defining expenses into program services, general management, and fundraising categories (U.S. General Accounting Office, 2002).

Since the terrorist attacks on the United States that occurred on September 11, 2001, when an influx of donations were made confirming public support of and trust in nonprofit organizations, a significant shift has occurred in how the public views organizational effectiveness, with performance expectations being at the very center of the shift (Light, 2002). During a time when donations, public trust, and media attention for nonprofit organizations were at an all-time high, how September 11, 2001, funds were distributed by a few organizations created immense public outrage. The outrage, coupled with public mistrust ensuing from corporate scandals such as Enron and WorldCom,

created an environment ripe for increased scrutiny and demand for accountability (Light, 2002) still affecting the nonprofit sector.

In 2007, more than 40 million individual donors, 400,000 corporate donors, and 70,000 independent, corporate, and community foundations gave \$306 billion to more than 1 million nonprofit organizations (O’Flanagan, Harold, & Brest, 2008). Yet a 2007 Harris Interactive poll indicated only 10% of Americans strongly believe charities are honest and ethical (Hoffman, 2006) with donated funds, which some experts believe is a result of measuring and comparing financial efficiency, a misleading and faulted measurement (Pallotta, 2008; Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004a; Wing, 2009). And because “program spending and fundraising efficiency ratios are not useful indicators of fraud or dishonesty, the focus on these ratios shifts the gaze of regulators and commentators from dishonest charities to those who are disadvantaged in the market for lower costs” (Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004b, p. 4).

Financial efficiency is widely acknowledged as an imperfect measure of nonprofit organizational effectiveness. The GAO, IRS, charity watchdogs, and experts have cautioned using such data because of accounting discrepancies, the range of organizational demographics that affect efficiency ratios, and the oversimplified nature of the question in relation to the complex nature of the nonprofit sector (U.S. General Accounting Office, 2002). Given this general knowledge of the flaws, financial efficiency remains the most prevalent measure for nonprofit evaluation. A 2001 survey showed more than half of respondents said that the amount an organization spends on

administration and fundraising is the most critical factor when determining recipients for their donations (Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004b). Only 6% said the most important reason to give is the organization fulfilling its mission and making a difference (Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004b).

**Efforts to measure nonprofit organizational performance.** Although financial indicators provide a basis for comparison across the sector, a growing body of literature denotes these are not true indicators of organizational outcomes or effectiveness. Alternative measurement models have been developed and tested; however, most are designed for a specific organization type, location, or size. Some theorists have even purported that creating a performance measurement tool usable across the nonprofit sector is likely impossible (Herman & Renz, 2008). Many of the theorized and tested models focus on characteristics, practices, and factors that fall into overlapping categories but are limited by organizational demographics (Herman & Renz, 2008).

A model that could address the weaknesses of current models while adopting their strengths, and adapted for broad application across nonprofit organizational types and missions, would be an important tool for internal and external performance measurement. Lack of consensus within the field has prevented the development and wide acceptance of such a model in the nonprofit sector. Another point of contention in the debate is that the focus is on *how* to measure performance, or the methodology of performance evaluation, rather than *why* nonprofit organizational performance should be measured in the first place and by *what* criteria nonprofit performance should be measured.

Many concepts exist for how to define and measure nonprofit performance and effectiveness. The dissent about to whom nonprofit organizations should be accountable, for what reasons, and how to evaluate them has generated many conceptual and theoretical frameworks but little agreement among them (Benjamin, 2008). The centrality of organizational effectiveness in nonprofit organizational performance is generally agreed upon; however, little agreement exists on the definition, measurement, and constitution of effectiveness (Lecy et al., n.d.). Each new construction of nonprofit performance affects how resources are provided and expectations are contrived, which leaves the nonprofit sector and its leaders in a continual tug-of-war between whose definition and method to implement and follow.

As the number of nonprofit organizations increases, demands for accountability and proven results become even more important. There is widespread desire for an easily digested and understood benchmark by which to compare organizations, and financial efficiency is the only measure to satisfy the public to date (Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004b). Despite the efforts of researchers and experts to develop a standardized model for analyzing nonprofit effectiveness, many leaders of nonprofit organizations do not assess the organizational performance or focus on evaluation standards as demanded by external constituencies (Carman & Fredericks, 2010).

Lack of internal measurement, diversity in nonprofit missions and operations, and the vast range in size and age of organizations make broad application of an external measurement method difficult (Ebrahim, 2005; Herman & Renz, 1997; Lecy et al., n.d.). Funders and external stakeholders increasingly demand a demonstration of effectiveness



through valid, understandable performance metrics that are replicable and widely applicable (Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004b).

### **Problem Statement**

Current evaluation and assessment metrics often hurt rather than help nonprofit leaders improve organizational performance (Ebrahim, 2005) and provide external parties with a limited or incomplete understanding of outcome delivery (U.S. General Accounting Office, 2002). The importance of performance measurement is often neglected or downgraded in many nonprofit organizations because of insufficient resources and training. Competing demands from external parties for a wide range of information are often unrelated to the mission or performance improvement (Baruch & Ramalho, 2006; LeRoux, 2009; Mayne, 2007), which is further compounded by the ever-changing socially constructed definitions of performance expectations. Although there is ample discourse regarding the need for changes and current replacement models, there is little consensus within the field on why, what, and how to evaluate while maximizing nonprofit organizational performance, outcomes, and mission achievement.

The general problem is that many theoretical models exist, offering different perspectives on nonprofit organizational performance measurement; however, few have been tested, have been agreed upon, or have received notable recognition or adoption within the sector. The specific problem is a lack of understanding of the subjective opinions held by nonprofit leaders regarding the purpose, process, and content of nonprofit organizational performance measurement creates difficulty in designing a widely applicable model that addresses differences in opinion, which may be the reason

for the development of multiple unimplemented models. The difference in opinions must be identified before any consensus or compromise can be achieved in the creation of a new performance measurement model to guide and assess the sector's performance.

This Q methodology study explored nonprofit leader perceptions of effective performance measurement and the purpose, content, and process elements necessary to create an effective model to measure nonprofit organizational performance. A sample of 22 nonprofit leaders in the United States were surveyed to gather, analyze, and document their perspectives regarding effective organizational performance measurement. More specifically, nonprofit leaders with experience in nonprofit management and performance measurement were targeted.

For the purposes of this study, nonprofit leaders were defined as (a) academic experts who have published within the field in the past 5 years or (b) professionals with 5 or more years of leadership experience within the field. Professionals are nonprofit organizational senior management, evaluators, grantors and funders, consultants, and watchdog agency leaders. Members of nonprofit organizational management and performance measurement groups or associations that focus on improving the nonprofit sector were identified as well. The goal of the research was to determine the characteristics of an effective organizational performance measurement model that could alleviate issues associated with current practices.

### **Purpose Statement**

The purpose of this Q methodology study was to identify the subjective perspectives of nonprofit leaders in the United States concerning effective organizational performance measurement practices. Q methodology involves systematic analysis of

human subjectivity about a topic of study (McKeown & Thomas, 2013). This study attempted to identify patterns of beliefs among nonprofit leaders. A chain referral sample of nonprofit leaders who have experience in nonprofit management and performance measurement served to help clarify the vast amount of information related to nonprofit organizational performance measurement. From their subjective perspectives, participants identified the most important purpose, content, and process elements of an effective nonprofit organizational performance measurement model. Much like snowball sampling, chain referral sampling relies upon individuals with experience with the phenomenon of interest to provide a series of participant referrals; however, chain referral sampling provides strategic access to multiple networks (Penrod, Preston, Cain, & Starks, 2003). The study simultaneously involved three elements of performance measurement: (a) purpose (why organizations are measured), (b) process (how organizations are measured), and (c) content (what is measured). A single condition of instruction was used to determine whether the nonprofit leaders believed the elements were equally important to effective performance measurement or whether certain elements were deemed more or less important than others.

The intent of Q methodology is to uncover social perspectives about the topic of study. Q methodology swaps the focus of the survey method, viewing statements as subjects and participants as variables (Sickler et al., 2006). Participants gave subjective meaning to the Q-set as they sort the sample of statements about performance measurement characteristics, which were correlated between participant profiles for similar viewpoints (van Exel & de Graaf, 2005). Data were analyzed to determine whether different factors from the collective of the individual perspectives exist that are

discernible only to nonprofit leaders experiencing the phenomenon under investigation (Stephenson, 1953). The objective of this study was to document nonprofit leader perceptions of effective performance measurement, which could help in the creation of a widely accepted external performance measurement model, thereby replacing financial efficiency in measuring nonprofit organization performance.

### **Significance**

Researchers have debated the concept of organizational effectiveness for many decades, with different waves in thought, definitions, and proposed evaluation methods. The debate has become more relevant in the nonprofit sector in the last decade as demands for accountability and return on social investments have increased. The following sections contain a description of the significance of the study in general and as applied to nonprofit organizational leadership.

**Significance of the study to the field.** As the nonprofit sector continues to grow in size, and competition for declining resources becomes fiercer, external parties demand increased accountability and proven results from nonprofit organizations (Ebrahim, 2005). Without an established definition of organizational effectiveness or method for performance measurement, a widening concern is that the media and charity watchdogs are encouraging “donors to make decisions based on simplistic measures that can never substitute for real assessments of whether a nonprofit group makes a difference” (Wing, 2009, p. 13).

Identifying nonprofit leader perceptions regarding performance measurement helped to document the subjectivities that could prevent or promote a widespread adoption of new performance measurement practices. A model that measures nonprofit

effectiveness, performance, and impact is necessary to replace financial efficiency standards. The identification of statistically significant social perspectives regarding the characteristics of effective performance measurement could lead to the development of a more beneficial performance measurement model for external parties and nonprofit organizations.

Perception is the “process by which people select, organize, interpret, retrieve, and respond to information from the world around them” (Schermerhorn, Hunt, & Osborn, 2008, p. 80). Feelings and actions are responses derived from people processing informational inputs through perception (Schermerhorn et al., 2008). Perception is a filter affecting individual responses, including feeling, thinking, and action. An individual’s perceptions, therefore, greatly influence his or her reactions and responses to situations (Schermerhorn et al., 2008). How nonprofit leaders perceive performance measurement affects their decision whether to conduct performance measurement or not, the importance they place on this management function, and other decisions related to organizational performance.

Individuals have deep-rooted mental models composed of internal images of how the world works (Senge, 1990; Starkey, Tempest, & McKinlay, 2004). These mental models can have a powerful influence on individual action because of their effect on how the individual perceives the world. Understanding nonprofit leader perceptions of performance measurement may help to uncover challenges in achieving a wide-scale replacement for financial efficiency in the evaluation of nonprofit organizations. Individual perception affects individual learning and decision-making processes, which are the foundation of organizational action and learning (Starkey et al., 2004).

Understanding nonprofit leader perceptions might help to uncover how nonprofit organizations are positioned for organizational learning and making performance improvements.

External constituencies would benefit from a performance measurement model that presents an accurate picture of organizational performance and social outcomes, and facilitates the benchmarking or comparison of nonprofit organizations. Such a tool could educate both donors and nonprofit leadership about which characteristics, practices, and behaviors position the sector for effectiveness and maximum community impact. The identified factors could serve as a guide for more beneficial and empowering internal performance improvements by nonprofit leaders.

A change in performance measurement and evaluation methods could help to eliminate donor and regulator expectations and behaviors that hinder rather than help nonprofit organizations from accomplishing their missions. Because of nonprofit organizational resource dependencies and the institutionalization of external expectations, if external constituencies change the performance measurement questions they ask, nonprofit organizations will look inward and change their focus. Ultimately, nonprofit organizations will be better positioned to affect true social change, thereby eradicating the problems they were established to solve. “To a large extent, the way we measure success determines the success we will achieve. Unmeasured things cannot be easily replicated, or managed, or appreciated” (Spitzer, 2007, p. 11). Changing the socially constructed definitions of organizational effectiveness and performance expectations in addition to altering what is measured can position the sector for increased social benefits,

use resources more efficiently, and solve the problems the organizations were created to address.

**Significance of the study to leadership.** Nonprofit leaders are increasingly subjected to changing and inconsistent socially constructed definitions for measurement and behavior (Herman & Renz, 1997). Because of resource dependence, leaders of nonprofit organizations often conform to and institutionalize the demands and needs of external parties rather than focusing on methods to accomplish their missions better (Froelich, 1999). The demands of an organization's donors pull nonprofit leaders off course, and rarely in directions that help leaders improve organizational learning and performance as opposed to meeting the needs of the donor (Lowell, Silverman, & Taliento, 2001). Externally imposed performance measurement usually fails to gain buy-in from nonprofit leaders and essentially reiterates the internal cultural resistance to evaluation (Spitzer, 2007).

One of the most underappreciated leadership roles is measurement leadership (Spitzer, 2007). If nonprofit leaders fail to make performance measurement a specific and intentional practice, "measurement in organizations will continue to become even more *dis-integrated, sub-optimizing, and non-transformational*. At the root of the problem are two misconceptions: 1) that measurement is primarily a technical activity, and 2) that no one individual has *overall* responsibility" (Spitzer, 2007, p. 121). Nonprofit leaders must take an active role in creating a culture of performance measurement and performance improvements if they are to advance the mission of the organizations they are entrusted to lead. Performance measurement is one of the most critical practices of an organization to determine and expose value creation and

destruction, and as such, organizational leaders must make measurement leadership a fundamental part of their jobs (Spitzer, 2007).

Documenting the purpose, process, and content elements that nonprofit leaders deem most important to nonprofit performance measurement might help to develop nonprofit evaluation methods that end these counterintuitive and counterproductive practices. Development of a performance measurement method that is feedback rich, improvement-based rather than punitive, and focused on the needs of the nonprofit organizations will help create buy-in among nonprofit leaders. Such a performance measurement method may help improve the effect and social benefit that nonprofit organizations provide for society.

### **Theoretical Framework**

Nonprofit organizations are social systems dependent upon external resources for survival (Scott & Davis, 2007). External influence and pressure threaten autonomy and demand conformity to and institutionalization of external standards and norms (Worth, 2009). The standards and definitions of effectiveness, impact, and performance are dynamic, subjective social constructions (Herman & Renz, 1997), which help to explain the ongoing dialogue for how to evaluate nonprofit organizational performance most effectively.

**Systems theory.** Von Bertalanffy and other early systems theorists pursued a general systems theory that could help explain the systems found in all fields of science (Denhardt, Denhardt, & Aristigueta, 2009). The primary characteristic of an open system is the interdependence of an organization on its environment (Scott & Davis, 2007). The focus of this paradigm is the complexity and variability inherent within the environment



and the interconnectivity linking the network of environmental actors. By definition, nonprofit organizations are open systems because of the complex interconnectivity between organizations, their environment, and their stakeholders (Scott & Davis, 2007). Without donors, volunteers, and communities, nonprofit organizational existence would be impossible and negated as these parties provide critical resources while receiving benefits from the social objectives of nonprofit organizations.

In an open system, an organization will function by acquiring and transforming external inputs into the output provided back to the environment (Starnes, 2000). On the most basic level, a nonprofit organization receives donations of time and resources and transforms these inputs into a service rendered to the community. A danger to open systems is entropy; the organization will fail without continued inputs of energy in the form of resources, people, ideas, and so forth (Katz & Kahn, 1978). Without positive support from the environment, such as continued tax-exempt status or donations of time and money, the nonprofit organization will be degraded and eventually fail. A nonprofit organization is inherently dependent upon the environment for existence and creates a network of nonprofit organizations subject and often obedient to external pressure.

**Resource dependence and institutional theories.** Resource dependence theory helps explain the external demand for accountability and performance measurement in the nonprofit sector (Worth, 2009). Organizations are dependent upon their external environment, constrained by their need for resources, and can lose autonomy (Froelich, 1999; Worth, 2009). Without managing resource dependence through strategic resource acquisition, nonprofit organizations risk isomorphism as they adapt to the constraints of their environment (Worth, 2009). Institutional theory accounts for nonprofit

organizations adopting external standards and expectations and the tendency for organizations to assimilate and adhere to external standards and norms (Worth, 2009). Dependence creates the need to demonstrate responsiveness and incentive to devote more time and attention to some stakeholders rather than to others, especially when catering to an external governance agenda that often consumes a disproportionate amount of time and resources (LeRoux, 2009).

Because outcomes and outputs are difficult to quantify in the nonprofit sector, organizations endure strong pressure to conform to behavioral expectations (Callen, Klein, & Tinkleman, 2003). Conformance to external pressure and the need for legitimacy explain nonprofit organizational adoption of performance measurement and evaluation (Carman, 2010). Incompatible stakeholder expectations increase complexity and the difficulty of internal decision making and organizational performance (Judge, 1994; Whetten, 1978), coupled with external pressure from individual, corporate, and government donors, which affects nonprofit behavior when resource dependence is high or diversification of funding is low (Froelich, 1999).

A successful organization diversifies resource providers to minimize dependence, reduce external control, and maintain autonomy (Froelich, 1999). At an optimal state, an open system organization will change, grow, and adapt to the demands of the environment; without continued input from and interaction with the environment, the organization will deteriorate. These processes are inherent to open systems, yet the socially constructed nonprofit ideology creates a tension between the nature of the organization and the constraints placed upon it both internally and externally (Pallotta, 2008). Performance expectations and organizational effectiveness are externally driven

social constructions adopted by nonprofit organizations because of resource dependence and institutionalization of the expectations of those they depend upon.

**Social constructionist theory.** An ontological perspective, social constructionism considers human actions, beliefs, and knowledge as creators of reality (Herman & Renz, 1997). Individuals creating reality might reach states of consensus or disagreement about the meaning assigned to social phenomena. Organizational effectiveness, performance measures, and social impact are social constructions defined by individual interaction and knowledge. Subjectivity influences the conceptualization of organizational effectiveness, which is not a stable or definite construct as new interactions influence the individual's understanding and definition of the construct.

Pallotta (2008) explained that the nonprofit sector is hindered by a dated, restrictive, and counterproductive nonprofit ideology. The focus of this socially constructed ideology is not on organizational outcomes, but rather on the methods employed by the organization. There is also the desire for immediate gratification or movement rather than a focus on long-term results. The system created from this ideology prevents effective organizational results and the eradication of those problems nonprofit organizations were established to solve. The accepted nonprofit ideology based effectiveness upon financial efficiency, provided a simple basis of comparison for the nonprofit sector, and was the best at the time. The old nonprofit ideology will no longer serve because of exponential growth in the nonprofit sector, the increase of information available resulting from advances in technology, and a growing demand for accountability.

A new paradigm is necessary for nonprofit performance expectations, effectiveness, and a method for assessing and evaluating organizational performance that allows nonprofit organizations to succeed. Despite the proposal of many ideas and methods, few have gained support or widespread adoption as the new paradigm for nonprofit performance evaluation. To gain legitimacy, prove accountability, and maintain positive relationships, nonprofit organizations must conform and institutionalize externally imposed standards, whether they are in the best interest of the organization's mission achievement or not. At the heart of the problem is the lack of an agreed upon definition of effectiveness and method for performance measurement. Effectiveness is subject to how the vast numbers of stakeholders define organizational effectiveness; no reigning body is offering a concrete definition to measure against. Thus, organizational effectiveness is what stakeholders or multiple constituencies determine it to be (Herman & Renz, 2004).

Existence as open, resource dependent systems, reliant upon interactions and support from outside agents, leads to operational difficulties for nonprofit organizations. Nonprofit organizations must interact with their environment to flourish, and without external funding sources to provide revenue, very few would survive. Social constructions of nonprofit performance, in addition to the priorities of external funding and regulatory sources, create an environment where nonprofit organizations are subject to external standards and expectations that often directly contradict their ability to be effective. In the name of assistance, many nonprofit organizations experience increased difficulty in mission attainment.

An existence as open systems subjects nonprofit organizations to a vast pool of interested stakeholders, all with their own opinions and beliefs about nonprofit performance and effectiveness. Resource dependencies affect which opinions and beliefs nonprofit leaders hear most and which socially constructed performance expectations and standards nonprofit organizations institutionalize. The discourse will be drawn from this theoretical framework, specifically, the discourse occurring between the network of interconnected nonprofit organizational stakeholders about the socially constructed beliefs and opinions held about nonprofit performance and effectiveness. How performance and effectiveness are most often shaped by the resources stakeholders offer nonprofit organizations and the pressures experienced by nonprofit leaders to institutionalize these expectations will be examined.

### **Research Questions**

The widespread dialogue regarding performance measurement can be reduced to three common elements: (a) purpose (why organizations are measured), (b) process (how organizations are measured), and (c) content (what is measured). Q methodology provides a systematic approach to understanding human subjectivity (McKeown & Thomas, 2013) and permits nonprofit leader perceptions of effective organizational performance measurement characteristics to be identified. The research questions for this Q methodology study are as follows:

RQ1: What do nonprofit leaders believe to be the most important elements of effective nonprofit organizational performance measurement?

RQ2: What are the characteristics, behaviors, or practices nonprofit leaders believe are critical to developing an effective nonprofit organizational performance measurement model?

Many different conceptions exist for how to define and measure performance and effectiveness. The dissent about whom nonprofit organizations should be accountable to, for what reasons, and how to measure them appears in many conceptual frameworks, with few in agreement (Benjamin, 2008). The centrality of organizational effectiveness is generally agreed upon; however, there is little agreement on the definition, measurement, and constitution of effectiveness (Lecy et al., n.d.). Each new construction of nonprofit organizational performance affects how resources are provided and expectations are contrived, which leaves the nonprofit sector and its leaders in a continual tug-of-war between whose definition and method should be implemented and followed.

Many ideas for how to evaluate nonprofit organizations exist, but there are relatively few if any large-scale empirical studies that involve examining the evaluation practices of nonprofit organizations, and most studies are narrow in focus or are case studies examining the practices at a few organizations (Carman, 2009). More theoretical and conceptual frameworks have been offered than empirically tested. With the rich field of opinions and methods, it is not surprising that the empirical findings demonstrate the difficulty nonprofit leaders have in designing and implementing effective performance measurement systems, for what purpose, and collecting what information.

### **Nature of the Study**

A researcher can use many methodologies to study nonprofit leader perceptions of performance measurement; however, the method employed should be dictated by the

research problem and research questions (Creswell, 2005). The selection of research design includes a set of decisions made by the researcher about the topic under study, among which are population, which research methods to employ, and for what purpose (Babbie, 2008), to provide the plan for the data collection, measurement, and analysis (Cooper & Schindler, 2008). In nonexperimental research, research designs can be causal, descriptive, or exploratory (Salkind, 2008). An exploratory design is appropriate for this study because the goal of the research is to identify nonprofit leader perceptions regarding performance measurement rather than to explain or establish causal relationships.

A critical evaluation of different research methods resulted in the determination that a mixed-method research design, and more specifically a Q methodology design, would be superior to other quantitative or qualitative methods. Q methodology includes the strengths of qualitative and quantitative research methods and, through the systematic examination of subjectivity, may reveal connections overlooked by other techniques (Akhtar-Danesh, Baumann, & Cordingley, 2008). Q methodology “brings out patterns or clusters of views and attitudes held by certain groups of people, as expressed in everyday or accessible communication” (Wolf, 2010, p. 255).

Q methodology allowed for an exploratory study of the subjective perceptions held by nonprofit leaders regarding nonprofit performance measurement. Subjectivity is fundamentally an individual’s point of view and may be in opposition to typical research resonating from the external standpoint of the researcher (S. R. Brown, 1980). Q methodology supports the exploration of the participant’s internal standpoint and “is used to discover typologies, to discern differences in opinion groups” by identifying common

factors (Peck, 2010, p. 4). The factors essentially function as “hypothetical individuals” (Peck, 2010, p. 4) and provide a window into the beliefs about effective performance measurement.

The focus of Q methodology is “on the individual rather than the collective and preserves the integrity of the individual in the final result. Normative methods such as sample surveys . . . deal with the collective behavior of groups . . . but in the process individual identity is lost” (Peck, 2010, p. 1). Q methodology transforms subjective statements into operant, objective outcomes in a feasible and effective manner (Akhtar-Danesh et al., 2008). Although critics of Q methodology raise concerns over representativeness, the intent of the methodology and this study was not to identify what portion of the population holds a particular view, but rather to uncover the views held among nonprofit leaders, and Q methodology is the ideal method to gather and analyze this information. Therefore, representativeness is a matter of inclusion of the population of all viewpoints rather than of individuals.

In the United States, there are 1.5 million tax-exempt 501(c)3 nonprofit organizations (Independent Sector, 2012a), each led by an executive director, chief executive, founder, or other leader, in addition to an unknown number of experts, researchers, evaluators, professionals, and consultants participating in the nonprofit organizational management and performance measurement dialogue. It is from this target population that participants were recruited. As it would be cost and time prohibitive to study the entire population, the participant recruitment was limited to a portion of the population (Cooper & Schindler, 2008).



This study collected the perspectives held by 22 nonprofit leaders, targeting those with experience in nonprofit management and performance measurement within the United States as the primary source of data. A relatively small number of participants are necessary because they are the means for measuring variables rather than representing a population (Sickler et al., 2006). A 3:1 ratio of statements to participants is most common, with the lowest ratio of statements to participants being 2:1 (Webler, Danielson, & Tuler, 2009). There must be enough participants to have four or five to define each factor, but additional participants have little impact on the factor scores because Q methodology results are generalizable to the defined factor type rather than the population (S. R. Brown, 1980).

A growing body of Q studies has affirmed that two to five social perspectives about a given topic are typically identified, and only four or five participants are necessary to define a factor (perspective; Webler et al., 2009). Identified through chain referral sampling, participants in this study were specifically asked to refer those in the field with a known difference in opinion from their own in an attempt to reduce bias and increase the researcher's ability to find different social perspectives held by nonprofit leaders.

Participants comprising the P-set should not be selected at random, but rather have a theoretical relevance to the study and a distinctive stance on the topic that will help define a factor (S. R. Brown, 1980). For the purposes of the study, nonprofit leaders were defined as (a) academic experts who have published within the field in the past 5 years or (b) professionals with 5 or more years of leadership experience within the field of nonprofit management and performance measurement.

Nonprofit leaders can help clarify the vast amount of information related to nonprofit organizational performance measurement and identify the most important purpose, content, and process elements to yield valuable data for external parties and internal performance improvement alike. The information was clarified by participants through the Q sort, a process in which participants sorted a collection of 42 Q sample statements representing the spectrum of opinions about nonprofit organizational performance measurement. The Q sample is a representative list of statements derived from the concourse that has been condensed to reduce ambiguity and duplication yet still represents the main ideas, opinions, and views collected in the concourse, typically between 40 and 50 statements (Wecht & Rago, 2006). During a Q sort, participants rank statements in a significant order with respect to their point of view (S. R. Brown, 1980) under a forced-choice distribution according to the sorting directive or condition of instruction. The Q sort is a reflection of the participant's thinking, evaluation, and interpretation of the Q sample stimuli (S. R. Brown, 1980).

Participants received Q sort materials through the U.S. mail. Each packet of materials consisted of a standardized list of instructions for completing the Q sort, a list of 42 alphabetized statements, and a short demographic questionnaire (see Appendices C, D, and E). Upon completion of the Q sort and subsequent recording procedures, participants returned Q sort materials in the provided self-addressed stamped envelope through the U.S. mail to the researcher within 2 weeks of receiving the materials. Participants received a follow-up telephone call from the researcher after 2 weeks.

Although the sample is small, is nonrandom, and may not be broadly applicable to individuals with different experiences (Akhtar-Danesh et al., 2008), the goal of a Q

methodology study is to identify different patterns of thought rather than the numerical distribution within a larger population (Akhtar-Danesh, Baxter, Valaitis, Stanyon, & Sproul, 2009). Correlation, factor analysis, and factor scores were used to analyze participants' Q sorts (McKeown & Thomas, 2013) to identify common performance measurement characteristics that nonprofit leaders believe are most important to the evaluation of nonprofit performance. Analysis uncovered patterns from individual responses that can be composed into social perspectives (Sickler et al., 2006). PQMethod 2.33 software facilitated analysis of Q sort results by computing the data (Schmolck, 2002).

### **Definition of Terms**

The academic setting and the Q methodology study included the application of many unique terms in relation to nonprofit performance measurement and organizational effectiveness. The following list of definitions includes commonly used words in this Q methodology study of nonprofit performance measurement.

*Concourse:* The concourse is a collection of statements relevant to the topic of study containing all related aspects of all discourse about the topic and comprising the raw material for Q methodology (van Exel & de Graaf, 2005).

*Financial efficiency:* Financial efficiency is the financial ratio analysis that determines the percentage of a nonprofit organization's budget spent on fundraising expenses, thereby exhibiting how efficient the organization is at raising funds (Judge, 1994).

*Nonprofit organization:* A nonprofit organization is an organization managed by a voluntary board of directors and afforded a special 501(c)3 tax-exempt status by the IRS

because it retains surplus funds to fulfill socially beneficial goals rather than to distribute funds to owners or shareholders (Independent Sector, 2012a).

*Organizational effectiveness:* According to Mensah, Lam, and Werner (2008), “While efficiency involves the quantitative ratio of inputs to outputs, effectiveness involves the softer concept of organizational goals versus organizational achievements” (p. 325).

*Organizational performance:* Lebas and Euske (2002) noted organizational performance involves “doing today what will lead to measured valued outcomes tomorrow” (p. 68).

*Overhead:* Overhead is the amount of money expended by a nonprofit organization on fundraising, management, and general administrative expenses (Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004a).

*Performance measurement:* Performance measurement is a management tool or process that utilizes a measurement system comprised of measures, a measurement process, and the technical infrastructure required to implement the assessment of defined performance expectations (Spitzer, 2007).

*P-sample (P-set):* A P-sample, or P-set) is a group of theoretically relevant participants in a Q methodology study who complete a Q sort (S. R. Brown, 1980).

*Q sample (Q-set):* The Q sample, also referred to as a Q-set, is a representative list of statements derived from the concourse and rank ordered by participants during the Q sort (Akhtar-Danesh et al., 2008).

*Q sort*: The instrument that facilitates a Q methodology study, the Q sort is the process by which participants rank order statements in the Q sample within the distribution based upon the relevant importance to the participants (Webler et al., 2009).

### **Assumptions**

In research, assumptions are the ideas or beliefs taken for granted without being tested for the conclusions being made (Neuman, 2006). This study included six assumptions. The first assumption was that nonprofit organizational performance should be measured. Performance measures and performance measurement “can have a *transformational impact* on the way people in organizations view their work, their products, and their customers” (Spitzer, 2007, p. 2). Performance measurement is not about the numbers but rather the perception, understanding, and insight an organization can gain from this important management tool, which organizational leaders of the future must make a regular practice to ensure survival (Spitzer, 2007).

The second assumption was that selected participants had the requisite leadership experience and knowledge of the nonprofit sector to understand the instructions for completing the Q survey adequately. Clear expectations of the desired leadership experience and knowledge are outlined as part of the chain referral sampling procedure to identify and include nonprofit leaders who meet these expectations. Each participant’s experience was confirmed by their postsort questionnaire responses. The third assumption was the survey tool was appropriate for collecting nonprofit leaders’ perceptions of nonprofit organizational performance measurement. The Q sort is the method by which information is collected for subsequent factor analysis. In essence, the Q sort is a reflection of the participant’s thinking, evaluation, and interpretation of the Q

sample stimuli (S. R. Brown, 1980) comprised of a range statements about nonprofit organizational performance measurement.

The fourth assumption was the nonprofit leaders and experts included in this study provided an appropriate representation of the opinions of nonprofit leaders in general about nonprofit performance measurement. It was also assumed that participants responded honestly and openly about their opinions of the topic and respond positively to participating in the study. Individuals who agreed to participate in the study received a request to read, sign, and return the informed consent form, which informs participants of confidentiality, the level of risk associated with the study, the nature of voluntary participation, and the ability to withdraw from participation. Return of the informed consent form and participation in this study revealed the participants' interest in the topic of study.

Assurance that all printed and electronic data and forms remained secure and locked, with only the researcher having access, should help elicit truthful answers from participants. Participants were assigned unique random identifiers rather than using their names or organizational affiliations to ensure confidentiality. "Confidentiality is maintained when anything that is learned about the participant is held in the strictest of confidence" (Salkind, 2008, p. 64). The anonymity of participants was ensured through the assignment of unique identifiers to each participant. A single master list of participants and identifiers was maintained (Salkind, 2008). The unique identifiers were used in all subsequent data collection and analysis documentation, thereby ensuring participant anonymity.

As in all Q methodology studies, the final assumption was that the lived experiences and perceptions of participants provided a reasonable foundation for all knowledge gained from the study. Q methodology allows for the exploration of the participants' internal standpoint (S. R. Brown, 1980) and taps into the subjective opinions, beliefs, and knowledge of the individual gained from perception and experience. "Q methodology is our window to learning about subjective reactions or responses to the issues confronting professionals daily" (Montgomery, 2010, p. 1). The intent of Q methodology is to uncover common social perspectives about the topic of study. Q methodology researchers collect "a small number of distinctly different individual perspectives and then employ non-parametric statistics to reveal the underlying social perspectives" (Sickler et al., 2006, p. 356) and in this study revealed social perspectives regarding nonprofit organizational performance measurement and potentially uncovered new theoretical ideas or positions worthy of further study. Participants were informed of their right to withdraw from participation at any point during the study for any reason of their choosing, including having inadequate knowledge and experience to complete the Q sort.

### **Scope**

The scope of this study was to use Q methodology to document the subjective perceptions nonprofit leaders hold for the purpose, process, and content elements of effective nonprofit organizational performance measurement. Participants were nonprofit leaders in the United States who have experience in nonprofit management and performance measurement. Initial contact with potential participants occurred through telephone, e-mail, or U.S. mail. Potential participants were identified during the

literature review and through research of charity watchdog agencies, foundations, and professional evaluators and nonprofit consultants. Contact was also made through the researcher's personal and professional network. For the purposes of this study, nonprofit leaders were defined as (a) academic experts who have published within the field in the past 5 years or (b) professionals with 5 or more years of leadership experience within the field. While each of the participants had a different role within the nonprofit sector, each had enough knowledge and experience with nonprofit organizational management and performance measurement to express an opinion about the most important characteristics of an ideal performance measurement model.

### **Limitations**

Limitations are the restrictions, limits, weaknesses, or problems identified by the researcher (Creswell, 2005). Q methodology "has the rather obvious disadvantage that it is designed to study subjectivities and thus is not suited to study matters of fact" (de Graaf & van Exel, 2008, p. 70). This characteristic makes Q methodology the ideal methodology for exploring participants' preferences, motives, tastes, feelings, and understanding of a topic, which are the underlying parts of personality that influence behavior (van Exel & de Graaf, 2005).

Although the sample was small, purposive, and may not be broadly applicable to individuals with different experiences (Akhtar-Danesh et al., 2008), the goal of this Q methodology study was to identify different patterns of thought rather than the numerical distribution within a larger population (Akhtar-Danesh et al., 2009). As such, findings are not generalizable to the larger population but to the concourse sampled for the Q-set. The definitions of validity and reliability are different in Q methodology because the



expressed perspective is derived from the participant's viewpoint rather than the researcher's external standpoint (S. R. Brown, 1980; Hathcoat & Montgomery, 2010). Because the unit of measurement is the importance of a statement to the subject, conventional reliability and validity are not central to Q methodology (Osterkiil, 2010). Validity is only an issue when participants genuinely fail to represent their own perspective (Choi, 2010; Hathcoat & Montgomery, 2010). Reliability is comprised of two parts: the test-retest coefficients from individual sorts and the schematic reliability or the replication of perspectives across multiple studies (Hathcoat & Montgomery, 2010).

Another criticism of Q methodology is the questionable reliability and the subjectivity of the researcher in relation to the process of statements chosen from literature and sorted by participants (Thomas & Baas, 1992). Strategic sampling helped reduce researcher bias in the Q-set creation, as statements were methodically chosen to represent the concourse both theoretically and inductively (Webler et al., 2009). Researchers in previous studies have concluded a limited number of unique viewpoints exist for a given topic (S. R. Brown, 1980), allowing for study replication. Replication, regardless of population, is a critical component for proving reliability (Buchman, 2009). Other researchers have found Q sort test-retest reliability to be .80 or higher, and content validity is assessed through a literature review and testing by a team of three to five domain experts (S. R. Brown, 1980; Valaitis, Akhtar-Danesh, Eva, Levinson, & Wainman, 2007).

Another limitation was the brief amount of time and resources dedicated to conducting the research may have limited the expressed opinions and perceptions of the participating nonprofit leaders to those who could be reached within time and fiscal

constraints. As it would have been cost and time prohibitive to study the entire population, the study population was limited to a portion of the population (Cooper & Schindler, 2008). Because “the objective in Q methodology is to be able to describe typical representations of different viewpoints rather than to find the proportion of individuals with specific viewpoints” (Akhtar-Danesh et al., 2008, p. 764), identification of the size of the population is unnecessary.

### **Delimitations**

The decision to include only nonprofit performance measurement experts in the United States limited the possibility of cultural, political, regulatory, social, and legal differences from nonprofit leaders from other countries introducing factors irrelevant to nonprofit organizational performance measurement in the United States. Limiting the study to nonprofit leaders with 5 or more years of experience in nonprofit management and performance measurement excluded the perceptions of leaders running nonprofit organizations such as middle management, program directors, and other employees who may not have met the experience requirement, may not have firsthand experience, and may not have an understanding of nonprofit organizational performance measurement. These individuals might have different perceptions of performance measurement than those represented by the participants included in this study. Both these design choices limited the ability to generalize the findings beyond the U.S. nonprofit sector and among individuals who are not experienced leaders in nonprofit management and performance measurement.

Social constructions of performance, performance expectations, and organizational effectiveness, in addition to external demands driven by resource

dependence and institutionalization, position nonprofit performance as dynamic and inherently longitudinal. Because performance expectations and definitions of nonprofit organizational effectiveness are dynamic social constructions that change over time, another delimitation of this study was the time dedicated to collecting the social perspectives of nonprofit leaders. The findings of this study provide a snapshot of nonprofit organizational performance measurement, performance expectations, and organizational effectiveness at the time of the study.

Despite delimitations in generalizing the results and the ever-changing dynamic conception of nonprofit organizational performance, the findings of this study are pertinent to documenting the perceptions held by nonprofit leaders with experience in nonprofit management and performance measurement regarding the most characteristics of an improved performance measurement model. Further studies can determine whether the factors identified in this study are exhaustive, what distribution shares these beliefs among the larger population of nonprofit leaders, and how nonprofit organizational performance measurement and organizational effectiveness have changed and continue to change over time.

### **Summary**

The nonprofit sector is growing and changing at a rapid pace, as are the external pressures for accountability and results. The focus of performance measurement practices is on measures that do not provide accurate or helpful information for internal performance improvements and external parties. Although there is ample discourse calling for new performance measurement models to replace financial efficiency, dissent has prevented the broad application of a new model in the sector.

As resource-dependent open systems, nonprofit organizations internalize the external expectations of resource providers, which have led to the adoption of harmful performance measures (Ebrahim, 2005; Froelich, 1999). Within the demands of daily operations, performance measurement is often underpracticed and seen as highly suspect in many nonprofit organizations. Clarifying the vast differences in opinions held by nonprofit leaders (Lecy et al., n.d.) might help to identify commonalities that could help build the consensus or compromise necessary for a broad application of a new performance measurement model.

This study included Q methodology, a qualitative method using quantitative strategies, as a methodology that draws upon the strengths of both quantitative and qualitative methods while mitigating the weaknesses of each. Q methodology is an appropriate method for an exploratory study of the subjective perceptions of nonprofit performance measurement held by 22 nonprofit leaders, as subjectivity is the primary element of focus in Q methodology. This study collected the perspectives held by nonprofit leaders who have experience in nonprofit management and performance measurement within the United States as the primary source of data. Although findings of this study are not generalizable to the larger population, the goal was to find the typologies of opinions rather than their distribution among the population. Through a review of the literature, Chapter 2 includes a foundation of the background knowledge required to provide an explanation of how the study fits into the historical, theoretical, and current body of research.

## Chapter 2

### Literature Review

The prevailing nonprofit ideology or paradigm defines organizational effectiveness through financial efficiency. Donors are trained to ask only one question: How much of my donation goes to the cause (Pallotta, 2008)? This singular focus on overhead percentages, or the amount of money diverted to operations rather than the cause, has given donors a simple metric to compare a complex sector. This flawed evaluation metric directs donor attention to an incomplete understanding of organizational effectiveness and leads nonprofit leaders to hide their need for operational inputs (Ebrahim, 2005). As a result, organizational infrastructure is secondary to programmatic spending and organizational effectiveness is threatened.

The nonprofit sector is different from the public and private sectors, in that organizational goals are typically intangible and more difficult to assess. Internal and external evaluation methods have been proposed and tested, yet few, if any, have been applicable across a broad range of organizational demographics and characteristics. Ample dialogue exists for how to evaluate nonprofit organizations, yet there is a lack of consensus on how to best evaluate so that it is relevant to external parties and assists nonprofit leaders with performance improvement. Evaluation as practiced by the nonprofit sector tends to be externally focused and retrospective and rarely informs changes in practice (Emerson, Wachowiz, & Chun, 2000).

The purpose of this Q methodology study was to gain a better understanding of the subjective perspectives of nonprofit leaders in the United States concerning effective organizational performance measurement practices. Q methodology provides for

systematic analysis of human subjectivity about a topic of study (McKeown & Thomas, 2013). The study attempted to identify patterns of beliefs among nonprofit leaders. A chain referral sample of nonprofit leaders who have experience in nonprofit performance measurement was used to help clarify the vast amount of information related to nonprofit organizational performance measurement and identify the most important purpose, content, and process elements to yield valuable data for external parties and internal performance improvement alike.

Through a review of the literature, Chapter 2 contains the foundation of background knowledge required to provide an explanation of how the study fits into the historical, theoretical, and current body of research. Chapter 2 includes an overview of the historical literature about organizational effectiveness, performance measurement and evaluation, and theories of nonprofit performance. The literature review encompasses internal, external, and mixed method evaluation techniques; their relevant strengths and weaknesses; and the absence of an agreed upon definition of nonprofit organizational effectiveness and performance expectations. Also addressed are an assessment of the current performance measurement and evaluation practices, the lack of consensus in the field of nonprofit organizational performance measurement, and a review of the purpose of nonprofit performance measurement.

Chapter 2 also includes a discussion regarding the internal and external characteristics that lead to effective performance measurement as well as analysis of the limited literature about nonprofit leader perceptions of organizational effectiveness and performance measurement. The chapter concludes with an examination of research methods and Q methodology, the method selected for this study. The objective of this

study was to document nonprofit leaders' perceptions of performance measurement, which could help in the creation of a widely accepted external performance measurement model, thereby going beyond financial efficiency in evaluating the performance of nonprofit organizations.

### **Documentation**

Scholarly books, peer-reviewed journals, and research documents were obtained through University of Phoenix and University of Alaska Anchorage databases, including EBSCOhost, ProQuest, ProQuest Digital Dissertations, JSTOR, Emerald, and SAGE Full-Text Collections. The research also included weekly scans of the websites of organizations committed to nonprofit organizational excellence, including Urban Institute Center of Nonprofits and Philanthropy, Indiana University Center on Philanthropy, Chronicle of Philanthropy, and the Hauser Center for Nonprofit Organizations at Harvard University. The literature review incorporated combinations of specific key word searches. Search strings included *nonprofit organizations*, *performance evaluation/measurement*, *performance metrics/models*, *external evaluation*, *outcome measurement*, *organizational effectiveness*, *accountability*, *capacity building*, *social impact*, *high impact [organizations]*, *leader perceptions*, and *Q methodology*. The search for primary sources was comprehensive though not conclusive because of the vast and dynamic characteristics of the online community. Given the waves of interest in the topic of organizational effectiveness and performance evaluation, research included sources from the 1990s through 2011, though earlier sources are included for a historical perspective of the topic. Although not all of the literature searched was applicable to this

study, Table 1 includes a summary of the literature accessed, reviewed, and cited in Chapter 2.

Table 1

*Summary of Literature Searched by Category (1990-2012)*

Category (by source type)	Accessed	Reviewed	Cited
Journals articles	3,844	476	83
Doctoral dissertations	41	13	2
Books	167	71	33
Other electronic sources	249	52	17
Totals	4,301	612	134

### **Historical Overview**

This section includes a review of various efforts to determine how to define and measure organizational effectiveness and performance among nonprofit organizations. Much of the literature includes overlapping definitions and models, some singular in focus and others including multidimensional constructs. For many decades, scholars have defined organizational success in terms of organizational effectiveness (Kondalkar, 2009); thus, an organization is successful if it is effective. The purpose of early organizational effectiveness research was to determine common metrics to compare the performance of organizations of different size and purpose (Lecy et al., n.d.).

One explanation for the increased demand to develop a metric for measuring organizational effectiveness among nonprofit organizations is the increased visibility of rating agencies and watchdogs among donors and the public in general (L. D. Brown, 2008; Ebrahim & Weisband, 2007; Unerman & O'Dwyer, 2006). The history of organizational effectiveness research has crossed multiple disciplines, including economics; accounting; operations; and personnel, organizational, and strategic management (Mensah et al., 2008). "While efficiency involves the quantitative ratio of



inputs to outputs, effectiveness involves the softer concept of organizational goals versus organizational achievements” (Mensah et al., 2008, p. 325). Measuring organizational efficiency is easy in comparison to organizational effectiveness, which often is not quantifiable or easily defined. An important point is that organizational effectiveness is not the opposite of organizational incompetence, but rather falls within a realm of organizational effectiveness, excellence, and competence (Ott & Shafritz, 1994).

Early phases of organizational effectiveness research involved an attempt to define dimensions of organizational effectiveness and create standardized terminology within the field, and the view of effectiveness was often as a one-dimensional construct. Researchers have used organizational goals, resource utilization, internal processes, and strategic constituencies as measures for organizational effectiveness. The goal attainment approach identified goals to measure performance, assuming that organizations are rational goal-seeking entities with clearly defined organizational goals (Price, 1968; Shilbury & Moore, 2006). If the organization’s goals are not specific and quantifiable, then they are not useful to this approach, and effectiveness cannot be measured.

The system resource approach roots effectiveness in an organization’s ability to attract and utilize resources, which ensures organizational viability (Shilbury & Moore, 2006; Yuchtman & Seashore, 1967). Although this approach measures organizational inputs and outputs, it does not examine organizational effectiveness holistically. The internal process approach involves examining the dynamic between employees, proposing factors such as trust, integrated systems, and smooth functioning as more accurate effectiveness measures (Shilbury & Moore, 2006; Steers, 1977). Although this approach involved examining important internal dynamics that affect organizational

effectiveness, the approach disregarded the external relationships and pressures faced by organizations. The strategic constituencies approach elicited key stakeholder definitions of organizational effectiveness, as stakeholders provide support and different interests in organizational performance (Connolly, Conlon, & Deutsch, 1980; Shilbury & Moore, 2006). Although this approach involved examining stakeholder perceptions of organizational effectiveness, a specific definition was unreachd and other critical factors were not examined.

The singular focus of each of these approaches and the ensuing criticism about them led to a conclusion that a more complex or multidimensional definition and approach to organizational effectiveness is needed (Lecy et al., n.d.). These models included multidimensional models (Cameron & Whetten, 1983; M. J. Foster & Lock, 1990; Zammuto, 1982), competing values models (Quinn & Rohrbaugh, 1983), contingency models (Lewin & Minton, 1986), and the balanced scorecard approach (Kaplan & Norton, 1996). These complex or multidimensional models often included or incorporated goal attainment, resource acquisition, and reputational dimensions to define or measure organizational effectiveness (Forbes, 1998; Lecy et al., n.d.). In response to a nonprofit's multiple constituencies, models to address multiple stakeholders emerged (Herman & Renz, 1997; Zammuto, 1984). Divisions within the field have led to more precise subfields (Lecy et al., n.d.), yet minimal effort has been made to integrate the divergent perspectives organizational effectiveness research has taken.

In a review of nonprofit organizational effectiveness literature from 1977 to 1997, Forbes (1998) found that, with the exception of one study, none of the researchers adopted a conventional organizational effectiveness approach but rather offered new

definitions or constructs. As a result of this literature review, a new, emergent approach to organizational effectiveness results, which shared three similar characteristics in which the meaning of effectiveness “is (a) created by the individual or organizational actors involved, (b) specific to the context in which it was created, and (c) capable of evolving as actors continue to act” (Forbes, 1998, p. 195). These studies gave critical attention to the role that perception and understanding have in organizational functions and the increasing consideration for the social construction of organizational effectiveness and organizational performance expectations.

Baruch and Ramalho (2006) reviewed a sample of 149 organizational effectiveness articles from 1992 to 2003 and found a variety of terminology issues and a lack of consensus for organizational effectiveness criteria, as most researchers used different operational definitions of organizational effectiveness and used terms interchangeably. Within the literature, a debate on organizational purpose occurred regarding whether the primary function was to fulfill goals (rational goal approach), to fulfill stakeholder expectations (strategic constituencies approach), or to fulfill multiple purposes (shift to multidimensional definitions). Some articles examined organizational effectiveness with separate constructs for private and nonprofit organizations, whereas others made a single construct for organizations with radically different functions. The sample demonstrated a lack of combined knowledge, as the body of literature did not build upon previous knowledge because researchers and theorists created new constructs rather than testing previously presented models. Most of the studies were one dimensional; nearly 41% focused on single or a few criteria, and despite multiple criteria, the studies were not multidimensional. The vast range in definitions resulted in a vague

understanding and definition of organizational effectiveness and an inability to measure organizational performance by a common standard.

Although organizational effectiveness has been presented as a way to measure and gauge organizational performance, the construct has shifted over time. Early researchers examined organizational effectiveness through one-dimensional constructs, though each was criticized for failing to address other important characteristics, factors, and dimensions of effectiveness. Although dialogue and literature suggest a move toward multidimensional organizational effectiveness criteria and constructs, audits of the literature have indicated that the definition and measurement of multiple dimensions remain a convoluted and difficult goal to attain. The increasing demand to measure nonprofit performance has maintained the relevance of organizational effectiveness within current dialogue and literature.

### **Theories of Nonprofit Performance**

Nonprofit organizations are social systems dependent upon external resources for survival (Scott & Davis, 2007). External influence and pressure threaten autonomy and demand conformity to and institutionalization of external standards and norms (Worth, 2009). These standards and definitions of effectiveness, impact, and performance are dynamic, subjective social constructions (Herman & Renz, 1997) that explain the ongoing dialogue for how to evaluate nonprofit performance most effectively.

**Systems theory.** During the 1960s and 1970s, systems theory emerged in social and organizational research and has become a systemic method for understanding the interconnectedness and consequences of human behavior (Denhardt et al., 2009). The primary characteristic of an open system is the interdependence of an organization on its

environment (Scott & Davis, 2007). Systems theory focuses on the complexity and variability of the parts, the interconnectivity linking the network of actors, and the flows among systems and the environment. By definition, nonprofit organizations are open systems because of the complex interconnectivity between organizations, their environment, and their stakeholders (Scott & Davis, 2007).

In an open system, an organization will function by acquiring and transforming external inputs into the output provided back to the environment (Starnes, 2000). On the most basic level, a nonprofit organization receives donations of time and resources and transforms these inputs into a service rendered to the community. External forces include general and task environments. The general environment includes indirect environmental influences, including legal, political, sociocultural, economic, and technological forces (Starnes, 2000). A change in the general environment can cause a shift in how a nonprofit organization will interact and function in the environment and can necessitate a change in the internal operations of the organization. The task environment directly affects an open system and consists of customers, competitors, and suppliers (Starnes, 2000). A nonprofit's task environment also includes donors, charity watchdogs, and governmental agencies, which influence management and strategic decisions made in the internal environment, comprised of the organization's employees and volunteers.

A danger to open systems is entropy; an organization will fail without continued inputs of energy in the form of resources, people, ideas, and so forth (Katz & Kahn, 1978). "Connections with 'external' elements can be more critical than those among 'internal' components; indeed, for many functions the distinction between organization and environment is revealed to be shifting, ambiguous, and arbitrary" (Scott & Davis,

2003, p. 31). Without positive support from the general and task environment, such as continued tax-exempt status or donations of time and money, nonprofit organizations will be degraded and eventually fail. A nonprofit organization is inherently dependent upon the environment for existence and creates a network of nonprofit organizations subject and often obedient to external pressure.

**Resource dependence and institutional theories.** Resource dependence theory helps explain the external demand for accountability and performance measurement in the nonprofit sector (Worth, 2009). Nonprofit organizations are not closed systems, and as such are dependent upon continued external exchanges (Heimovics, Herman, & Coughlin, 1993). Organizations are dependent upon their external environment, are constrained by their need for resources, and can lose autonomy (Froelich, 1999; Worth, 2009). Without managing resource dependence through strategic resource acquisition, nonprofit organizations risk isomorphism as they adapt to the constraints of their environment (Worth, 2009). Institutional theory accounts for nonprofit organizations adopting external standards and expectations and the tendency for organizations to assimilate and adhere to external standards and norms (Worth, 2009). “Institutional theory describes organizations’ efforts to gain legitimacy by embracing the norms, values, and mores prescribed by the environment in which they exist” (Worth, 2009, p. 49).

Because outcomes and outputs are difficult to quantify in the nonprofit sector, organizations endure strong pressure to conform to behavioral expectations (Callen et al., 2003). Conformance to external pressure and the need for legitimacy explain nonprofit organizations’ adoption of performance measurement and evaluation (Carman, 2010).

Incompatible stakeholder expectations increase the complexity of internal decision making and organizational performance (Judge, 1994; Whetten, 1978).

Nonprofit organizations' survival is dependent upon resource acquisition and maintenance ability (Pfeffer & Salancik, 1978). Scarcity and uncertainty in the environment complicate organizational existence. To acquire and maintain the resources necessary to function, leaders of successful organizations interact with individuals and groups who control resources. A danger is the organization can lose its autonomy. How reliant a nonprofit organization becomes depends upon the quantity, quality, and demand for the resources provided (Froelich, 1999). The fewer resources or providers, the higher the level of dependence the organization will experience. Organizational leaders must manage dependencies and avoid controlling demands through cooptation, acquiring countervailing power, and increasing resource providers (Froelich, 1999).

Dependence creates the need to demonstrate responsiveness and incentive to devote more time and attention to some stakeholders than others, especially when catering to an external governance agenda, which often consumes a disproportionate amount of time and resources (LeRoux, 2009). Resource dependence theory "forms the basis for a competing model of stakeholder management in which organizations can and do strategically place some stakeholder interests over others because financial performance (revenue growth) is contingent upon such a strategy" (LeRoux, 2009, p. 163). The source of external funding (corporate, government, or independent foundations) will transmit and embed different values into funded nonprofit organizational practices (LeRoux, 2009).

Froelich (1999) explained how external pressure from individual, corporate, and government donors affects nonprofit behavior when resource dependence is high or diversification of funding is low. Individual donors are highly unpredictable and unstable, creating a volatile dependence relationship. This relationship creates goal displacement, as nonprofit organizations modify behavior to satisfy donors, in addition to creaming, when nonprofit organizations shift programs to areas with greater donor appeal. As such, when the external environment pressures the nonprofit to concentrate on financial efficiency, the goal-directed behavior often shifts from mission completion to that of maintaining overhead ratios. Corporate donors increasingly donate out of self-interest, making decisions based upon marketing plans rather than benevolence. Self-interested corporate donor practices are compounded by the funding behavior of private foundations, as these donors are less likely to provide grants outside of their funding agendas, further influencing goal displacement. Government funds favor larger, uncontroversial recipients, many of whom receive more than half of their operating budgets from this source. While volatility decreases, organizations are subject to government-driven increases in professionalization, bureaucracy, and decreases in administrative autonomy.

Such targeted giving creates goal displacement and higher levels of formalization, which distracts attention from the mission and places it upon reporting activities (Froelich, 1999). Procedures, rather than outcomes, become the dominant focus for nonprofit attention and actions, creating a relationship that impinges upon mission attainment and service delivery. Leaders of a nonprofit will ideally determine which funding methods and resource providers best fit organizational goals and mission



delivery. Each have advantages and disadvantages, and ultimately organizational leaders must not be controlled by, but rather manage, resource dependencies (Froelich, 1999).

A successful organization diversifies resource providers to minimize dependence, external control, and maintain autonomy (Froelich, 1999). At an optimal state, an open system organization will change, grow, and adapt to the demands of the environment; without continued input from and interaction with the environment, the organization will deteriorate. These processes are inherent to open systems, yet the current nonprofit ideology creates a tension between the nature of the organization and the constraints placed upon it both internally and externally (Pallotta, 2008).

**Social constructionist theory.** An ontological perspective, social constructionism entails a belief that human actions, beliefs, and knowledge are creators of reality (Herman & Renz, 1997). Individuals creating reality may reach states of consensus or disagreement about the meaning assigned to social phenomena. Organizational effectiveness, performance outcomes, and social impact are social constructions defined by individual interaction and knowledge. “The social constructionist conception treats organizational effectiveness as stakeholder judgments formed in an ongoing process of sensemaking and implicit negotiation” (Herman & Renz, 1997, p. 188). Individual subjectivity influences the conceptualization of organizational effectiveness, which is not stable or static as new interactions influence how the individual understands and defines the construct.

Social constructionist theory posits that social processes derive from and sustain an individual’s knowledge. Knowledge, understanding of the world, and the meaning assigned to phenomena are social constructs. Three fundamental practices influence this

process: externalization, objectification, and internalization. Externalization occurs when individuals interact with the world, creating practices, processes, or artifacts from ideas (Burr, 1995). The ideas, once introduced to the social world, are repeated and become objects of consciousness, taking on elements of truth in their factual existence (Burr, 1995). These objects become natural, accepted parts of existence rather than connected to social interaction (Burr, 1995). Future generations enter a world in which these ideas exist and are internalized as a natural part of the world.

Lacan believed language was the tool that enables humans to connect with others, noting language creates the subject and the subject cannot exist without language (Sarup, 1993, p. 10). Each day when individuals communicate, they embed metaphors, meaning, and implications within their speech, intentionally or not. The language chosen by the individual contains meaning and subjective understanding. “Attention to the mediating role of language, interaction rituals, and categories will help explain how organizational routines and rules develop, stick, and fall into disuse” (Greenwood, Oliver, Suddaby, & Sahlin-Andersson, 2008, p. 278).

For centuries, humans have constructed theories and paradigms to gain an understanding and explanation of how the world operates. Paradigms provide explanation for phenomena, rules for practice or procedure, and standards for understanding the world. Kuhn (1996) explained while paradigms are accepted, they are not finite or infallible. In time, theories and paradigms that were standard explanation and practice have been changed, altered, and replaced by newer versions or radical changes as humans learn and evolve. Ideological revolutions often occur when one

paradigm explanation no longer serves those who use it. There is a need to adapt as the world continually changes.

Pallotta (2008) explained that the nonprofit sector is hindered by a dated, restrictive, and counterproductive nonprofit ideology. The basis of the ideology is in the Puritan ethic of deprivation and definition of charity, in response to the moral conflict between the individual gains of capitalism and the religious impetus for self-sacrifice (Pallotta, 2008). Charity allowed the wealthy to ease their conscience by elevating the needs of the poor. Charity was once the work of individuals, but over time, organizations were formed to carry out the work of the wealthy (Pallotta, 2008). The newly formed nonprofit organizations were inherently dependent upon their environment. Wealthy donors provided financial resources necessary to operate and the poor were the target for the services provided.

As part of this ideology, nonprofit organizations are not meant to prosper, but rather are forced to suffer and are ill equipped, as a deprived organization is morally sound (Pallotta, 2008). “The word ‘profit’ comes from the Latin noun *profectus* for ‘progress’ and the verb *proficere* for ‘to advance.’ Thus, the term ‘nonprofit’ means, literally, nonprogress. It is a dangerous unconscious statement of intent, or lack of it” (Pallotta, 2008, p. 3). From this foundation or guise of nonprogress, individuals in the nonprofit sector attempt to transform society. The focus of this ideology is not on organizational outcomes, but rather on the methods employed by the organization. There is also the desire for immediate gratification or movement rather than a focus on long-term results. The system created from this ideology prevents effective organizational results and the eradication of the problems nonprofit organizations were established to

solve. The ideology is also plagued with the belief that the nonprofit sector is not in direct competition with the for-profit sector. Both compete for discretionary dollars, yet the leaders of nonprofit organizations are prevented from using the tools of capitalism to serve their causes.

The widespread nonprofit ideology based effectiveness upon financial efficiency, provided a simple basis of comparison for the nonprofit sector, and was the best at the time. The old nonprofit ideology will no longer serve because of exponential growth in the nonprofit sector, the increase of information available resulting from advances in technology, and a growing demand for accountability. A new paradigm or method is necessary for assessing and evaluating organizational performance and allowing nonprofit organizations to succeed. Many ideas and methods have been proposed, yet few have gained support or widespread adoption as the new paradigm for nonprofit performance evaluation.

At the heart of the problem is the lack of an agreed upon definition of effectiveness and method for performance measurement. Effectiveness is subject to how the vast numbers of stakeholders define organizational effectiveness, as no one reigning body exists to offer a concrete definition to measure against. Organizational effectiveness is what stakeholders or multiple constituencies determine it to be (Herman & Renz, 2004). How the stakeholder defines organizational effectiveness will affect the rules, expectations, and demands placed upon the organization and embedded within institutional and resource support. The subjective perception of organizational effectiveness and performance measurement is at the center of creating a common understanding of how best to evaluate and assist nonprofit organizations.

L.D. Brown and Kalegaonkar (2002) explained the nonprofit sector is facing strong internal and external challenges. External challenges include public legitimacy; accountability; and relations with government, business, and internationally. Internal challenges include amateurism, restricted focus, material scarcity, fragmentation, and paternalism. Given these internal and external challenges, as resource dependencies and increased competition for resources affect nonprofit organizations, nonprofit organizations are increasingly subjected to socially constructed definitions of organizational effectiveness, performance measurement, behavioral norms, and expectations. To gain legitimacy, prove accountability, and maintain positive relationships, leaders of nonprofit organizations must conform and institutionalize these externally imposed standards whether they are in the best interest of the organization's mission achievement or not.

**A change in expectations.** An existence as open, resource-dependent systems reliant upon interactions and support from outside agents creates operational difficulties for nonprofit organizations. Nonprofit organizations must interact with their environment to flourish, and without outside funding sources to provide revenue, very few would survive. Social constructions of nonprofit performance, in addition to the priorities of external funding and regulatory sources, create an environment where nonprofit organizations are subject to external standards and expectations that often directly contradict their ability to be effective. In the name of assistance, many nonprofit organizations experience increased difficulty in mission attainment.

Nonprofit organizations cannot change into a closed system capable of providing their own resources. Current resource dependencies create an unhealthy power dynamic,

which leads to the institutionalization of harmful performance expectations and evaluation tactics. The only dynamic parts of the nonprofit sector are the socially constructed definitions of organizational excellence, performance, and performance expectations. Changing each of these to constructions that help the nonprofit organization improve performance and mission accomplishment will also change what beliefs, norms, and values the organization institutionalizes. For nonprofit organizations to improve at service delivery and solve the problems they were created to solve, change must happen.

### **Current Performance Measurement and Evaluation**

Performance is socially constructed and multidisciplinary, with each discipline offering its own definition and adding its own bias to the model (Holton, 1999). Menash et al. (2008) listed three main factors that dominate the definitions of organizational performance: (a) organizational resource-acquisition ability, (b) organizational goal attainment, and (c) organizational efficiency. “Performance management has been embraced by public and private funders to improve the effectiveness and efficiency of nonprofit programs” (S. R. Smith, 2010, p. 137), and as such it is an increasingly expected practice of nonprofit organizations.

In a review of the literature for performance measurement, Holton (1999) found that levels of performance measurement and performance indicators had different definitions in subsequent models and that bias clouded the subsystems in models as well. “It is difficult to practice performance improvement when there is no clarity about how to integrate the various ideas about performance dimensions and their measurement” (Holton, 1999, p. 28). Holton presented an integrated taxonomy of performance domains

that could be applied across disciplines and clarified performance measurement research. Four performance domains were presented: (a) mission, (b) processes, (c) critical performance subsystems, and (d) individuals. Within each domain are performance outcomes and performance drivers. Outcomes are units of effectiveness or efficiency produced by the system, subsystem, process, or individual (Holton, 1999; Kaplan & Norton, 1996). Performance drivers measure elements of performance that sustain or grow the performance dimension and are the leading indicators for future success (Holton, 1999; Kaplan & Norton, 1996).

Performance measurement and evaluation can be formative or summative. Formative evaluation occurs throughout the program or process or over time, whereas summative evaluation occurs at the end of the program, process, or period (Neuman, 2006). Formative evaluation is a structure for monitoring and generating continuous feedback (Neuman, 2006). Summative evaluation solely examines final outcomes and cannot help an organization improve throughout the process, but rather provides a snapshot of the final status (Neuman, 2006).

Most performance measurement processes and metrics take a summative rather than formative approach. Few evaluation efforts take a holistic view of performance measurement, but rather examine a fraction of organizational performance, likely because of the complexity associated with monitoring during the process in addition to the vast differences in organizational demographics and missions. The following sections outline current performance measurement and evaluation practices and delve into how each practice embodies a narrow view of performance measurement.

**Accountability and efficiency (external evaluation).** Accountability for public organizations includes three measures: (a) finances; (b) fairness toward employees, contractors, clients, and citizens; and (c) performance (Ospina, Diaz, & O’Sullivan, 2002). Traditional definitions of nonprofit accountability are typically less concrete, have prescribed standards for information disclosure and minimum behavior compliance, and offer nonprofit leaders no assistance in measuring organizational performance (Ospina et al., 2002). These accountability standards have more often been used to hold managers accountable for individual behaviors rather than organizational performance.

Accountability should be a relationship of processes and interactions that address questions of responsibility, discretion, reporting, reviewing, and revising (Whitaker, Altman-Sauer, & Henderson, 2004). “Procedures intended to provide accountability in relationships between governments and nonprofit organizations often focus on ways to catch and punish mistakes rather than on ways to improve service to the public” (Whitaker et al., 2004, p. 115). Accountability has an inextricable link to trust and transparency and implies enforcement, surveillance, responsibility, and exposure (Molnár, 2008).

Nonprofit organizational dependence on volunteers and philanthropy affords them special IRS tax-exempt 501(c)3 or 501(c)4 status. The organizations must file the IRS Form 990 annually if gross receipts are in excess of \$25,000 (Lampkin & Boris, 2002). Form 990 is the only comprehensive source of annual nonprofit organization financial information and includes revenue, expenses, net and total assets, balance sheets, program descriptions, board members, salaries of the top five employees, and lobbying information (Lampkin & Boris, 2002). The form has widely become a federal report for



charities from which external evaluators such as ratings agencies and charity watchdogs derive the information for their analysis and reporting. GuideStar digitizes more than 400 IRS Form 990 variables offering a comprehensive research database for public and private use (Easterly & Miesing, 2009). Although the form is the most widely used because it permits comparative analysis and longitudinal studies, questions regarding data accuracy, reliability, and gaps exist. Religious congregations and organizations with gross annual receipts under \$25,000 are not required to file the form, and further, nonprofit organizations with less than \$5,000 in gross annual receipts are exempt from registering with the IRS.

Escalation in the number of nonprofit scandals and concerns over unethical practices by nonprofit leaders has increased demand for accountability (Easterly & Miesing, 2009; Holt, 2006). Accountability is a substantial source of legitimacy in the eyes of the public for nonprofit organizations (Ospina et al., 2002). As part of the Sarbanes-Oxley requirement of increased accountability and transparency in governance, the IRS revised Form 990 (Holt, 2006). The new Form 990, released in 2007 and first filed in 2009, was designed to enhance mission transparency, financial information, operations, and resource use (Grant Thornton, 2008). The redesign was due to the demand for increased disclosure by the IRS to determine tax compliance and to focus on organizational governance, though not all new questions on the form are legally required to determine and verify tax status (Grant Thornton, 2008). The new form is intended to increase NPO accountability through a focus on policies, compensation, and joint ventures. Although some changes have been made, a 2007 GAO report identified nonprofit organizational internal governance in addition to administrative and reporting

requirements as policy issues for further examination and action (U.S. Government Accountability Office, 2007).

Financial efficiency is the most widely accepted and applied definition and evaluation method employed for assessing nonprofit organizational effectiveness (Judge, 1994). The prevailing logic was that the nonprofit organizational obligation of public service demands that effectiveness is determined by effective resource utilization and service outputs (Provan & Stewart, 1982). Other methods have not been adopted widely because of a complex nature or an inability to compare across the nonprofit sector. “Financial analysts have argued that relative spending on programs or overhead reflects ‘accountability’ and that relative costs of fundraising reflect ‘efficiency’” (Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004b, p. 1). Without other reliable or comparable data, charity watchdogs and the media use this analysis as a source of legitimacy (Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004b).

Three metrics are used to measure organizational efficiency: (a) ratio of administration expenses to total expenses, (b) ratio of fundraising expenses to total expenses, and (c) ratio of program expenses to total expenses (Callen et al., 2003). Better Business Bureau’s Wise Giving Alliance, American Institute of Philanthropy, and Charity Navigator offer external evaluation processes that rank or rate organizations, setting best practices and disclosure activities based on governance, financial performance, information transparency, and other factors (Sloan, 2009). Charity Navigator has developed profiles and ratings of nonprofit organizations solely on financial information from Form 990s (Urban Institute Center on Nonprofits and

Philanthropy & Indiana University Center on Philanthropy, 2004b). Charity Navigator uses a composite score based upon organizational efficiency and capacity, and organizations with the highest rating are expected to spend less than 15% on overhead. In December 2009, the chief executive officer (CEO) announced forthcoming changes, with the addition of accountability and eventually outcomes into the watchdog's evaluation practices (Berger, 2009). However, it is unknown how the change will affect evaluation or how organizational leaders will measure these new factors. The Better Business Bureau's Wise Giving Alliance, as part of the organization's accountability standards, established spending threshold ratio limits (Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004b). A Wise Giving Alliance pass rating, earned by spending more than 65% on programs (Better Business Bureau's Wise Giving Alliance, 2010), has provided positive financial returns for rated organizations (Sloan, 2009). The American Institute of Philanthropy rates nonprofit organizations on a letter grade scale (A through F) based upon the financial efficiency ratios and defines a satisfactory organization (or C rating) as one that spends at least 60% of funds on program-related expenses, and a highly efficient organization (A rating) as one that spends more than 75% of funds on program-related expenses (American Institute of Philanthropy, 2010).

In the absence of specific outcome data or outcome measurement, efficiency ratios have permitted comparative benchmarking, offering a strategy for program improvement and increased accountability (S. R. Smith, 2010). Ratings systems are normative frameworks, intended to influence nonprofit leaders and organizational behavior (Sloan, 2009). The standards for acceptable overhead expenses have been

widely adopted by the American public and nonprofit organizations alike. The problem, however, is that the acceptable limits promoted by charity watchdogs are not statistically proven and are arbitrarily set between 65 and 75% (Pallotta, 2008; Sloan, 2009; Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004a). “Minimizing overhead is a worthy goal, but it is counterproductive when taken so far that it stifles critical investments in the organization” (Lowell et al., 2001, p. 153). Charity watchdog weaknesses include an overreliance on poor quality data for simple analysis and ratios, excessive emphasis on efficiency rather than program effectiveness, and substandard qualitative analysis (Lowell, Trelstad, & Meehan, 2005).

Increases in efficiency do not equate to increases in organizational effectiveness, and vice versa (G. S. Smith, 1988). Use of a single performance indicator or dimension will lead to dysfunctional behavior to meet the expectation for the standard (G. S. Smith, 1988). Efficiency ratios favor older, larger, more established, and popular causes and discriminate against the newly established and unpopular or unknown causes whose ability to raise funds costs more (Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004b). The ratios inherently provide an incentive to NPOs that can minimize or hide overhead expenses and mask the true cost of NPO operations with creative accounting practices.

Empirically, Ritchie and Kolodinsky (2003) found six distinct financial performance measurement ratios from three performance-related categories: (a) fundraising efficiency, (b) public support, and (c) fiscal performance. Data were derived from Form 990s and factor analysis was applied. Ritchie and Kolodinsky proposed these financial ratios as a less expensive, multidimensional approach to measuring nonprofit

financial performance; however, financial management is the only dimension proposed with multiple financial criteria. Many known data quality issues exist with the Form 990, including data entry errors, incorrect organization types, inaccurate classifications, quickly outmoded data, infrequently updated files, limited financial information, incomplete staff and compensation data, and nonspecific data about complex organizations (Lampkin & Boris, 2002).

There is a wide discrepancy in how to organize expenses into program services, general management, and fundraising categories due to the discretion allowed by the IRS (U.S. General Accounting Office, 2002). As an example, the leaders of nearly 13% of nonprofit organizations reported no spending in management and general expense categories, an indication that all funds were spent on programs or fundraising, and it is understood that many organizations underreport their fundraising expenses, further indicating the majority of expenses are program related (Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004c). The situation is also complicated by the fact that Form 990 analysis completely disregards small nonprofit organizations that are exempt from filing, skewing the results and reporting an incomplete understanding of the sector (Toepler, 2003). Much of what is known or believed to be true about nonprofit organizations is derived from a fallible and incomplete data source and is focused on a single, financial dimension of performance.

In addition to the poor quality, inaccurate, and incomplete data from IRS Form 990s, the social goals of the nonprofit sector necessitate that organizational effectiveness have a broader definition than financial performance (Judge, 1994). Ebrahim (2005) explained the current accountability and efficiency evaluation standards focus attention

on external stakeholder demands, most typically those belonging to funders, rather than on organizational mission and outcomes. The focus is therefore on short-term output and efficiency rather than long-range goals and social change.

Typically, evaluation (a) occurs more often to appease funders; (b) is defined and framed differently depending on the source, which creates an inconsistent definition of organizational effectiveness; and (c) is rarely used to change behavior or inspire organizational learning (Ebrahim, 2005). Nonprofit organizations need simple, accessible evaluation methods, “particularly in a context where they are resource-dependent on funders and where outcome measurement is thus likely to serve as yet another system for grading performance rather than for improving it” (Ebrahim, 2005, p. 70). The greatest challenge is finding a balance between upward accountability to donors, meeting the needs of employees, and making positive change in organizational practices (Ebrahim, 2005, p. 73).

**Multidimensional evaluation.** Earlier waves of interest in organizational effectiveness research led to the belief that organizational effectiveness constructs must be multidimensional and have multicriteria. Although no single definition for organizational effectiveness exists, “the majority of authors agree that organizational effectiveness requires measuring multiple criteria and the evaluation of different organizational functions using different characteristics, and it should also consider both means (processes) and ends (outcomes)” (Shilbury & Moore, 2006, p. 9).

According to Lecy et al. (n.d.), a structured literature review of nonprofit organizational effectiveness articles published between the late 1990s and 2010 showed that most include three overlapping notions of organizational effectiveness based on (a)

goals, (b) resources, and (c) reputation. In their findings, few articles advanced the understanding of organizational effectiveness through empirical research and analysis; most articles were conceptual or theoretical, offering new assumptions rather than testing existing models. A majority of the articles failed to define organizational effectiveness adequately or to offer analysis at multiple levels, despite acknowledging the need to do so. Of the 64 articles sampled, only 12.5% were empirical, and the rest were theoretical, case studies, or program evaluations; only 43% classified methodology, and the rest were argumentative; and 67% did not define effectiveness, and those who did defined it quite differently. Lecy et al. concluded “a lack of empirical research, shared definitions, and sustained interdisciplinary efforts undermine progress in our collective understanding of organizational effectiveness and create obstacles to effectively mobilizing these literatures to answer current pressing questions about the future role of NPO/NGOs” (p. 3).

Research specifically examining nonprofit organizational effectiveness through multiple yet not necessarily overlapping dimensions of effectiveness includes project impact (Eisinger, 2002), financial efficiency (Mensah et al., 2008; Morgan, 2006; Ritchie & Kolodinsky, 2003; Tinkleman & Donabedian, 2007), managerial effectiveness (Lewis, 2001), board effectiveness (Herman & Renz, 1999), and effective use of partnerships and networks (Bacon, Kimble, & Taylor, 1989; Fredericksen & London, 2000).

Effectiveness is highly contextual, and the dominant approaches to nonprofit organizational effectiveness are reputational and hybrid multidimensional approaches (Lecy et al., n.d.). Lecy et al. (n.d.) presented four organizational effectiveness research domains derived from the current literature to help researchers and practitioners create a

composite of organizational effectiveness and eliminate the earlier fractioning of research. The domains were (a) organizational management, (b) program design and implementation, (c) responsiveness to the environment, and (d) networks and partnerships.

***Balanced scorecard.*** A multidimensional approach, Kaplan's (2001) balanced scorecard is a strategic planning tool for nonprofit leaders to integrate financial, programmatic, operational, and mission-related objectives. Although useful to strategic endeavors, the tool offers little in terms of outcome and impact measurement (S. R. Smith, 2010). An adaptation of the balanced scorecard, and to meet the characteristics of the nonprofit sector better, Moore (2003) introduced the public value paradigm, which is a multidimensional strategic planning tool consisting of a three-part triangle, including value, legitimacy and support, and operational capacity. As with the balanced scorecard, more attention is focused on the means (planning) than on the ends (outcomes), and defining value is problematic as it is dependent upon the collective public perception of value, which varies widely among citizens (Alford & Hughes, 2008).

***Multidimensional, integrated model for nonprofit organizational effectiveness (MIMNOE).*** Sowa, Seldon, and Sandfort (2004) presented the multidimensional, integrated model for nonprofit organizational effectiveness (MIMNOE). The MIMNOE prescribes characteristics of a measurement model that should incorporate two dimensions, management and program effectiveness, with two subcomponents, capacity and outcomes. Sowa et al. proposed a model necessitates both objective and perceptual models to capture organizational effectiveness dimensions, must allow for system structure variations, incorporates an analytical method with multiple levels of analysis,



and models interrelationships between the dimensions (Sowa et al., 2004). Although the MIMNOE provides a statistical model that has been useful for program evaluations, “it does not in itself make NPO effectiveness measurable and therefore does not offer much help in addressing true organization-level NPO effectiveness” (Herman & Renz, 2008, p. 408). MIMNOE’s definition of organizational effectiveness emphasizes program effectiveness heavily, which is not the same as organizational-level effectiveness.

***Logic models.*** Bailin (2003) explained that effective organizations have strong theories of change and have a developed, clear, communicable understanding of micro- and macrolevels of “what they do, with whom, at what cost; why that activity leads to measurable results; and how they define and recognize success when they see it” (p. 637). Typically, once identified, theories of change can be used to develop logic models. Logic models involve mapping the production process and focusing attention on better performance through resource, infrastructure, and program implementation rather than on specific, long-term outcomes (S. R. Smith, 2010). The Baldrige Award Criteria for Performance Excellence for nonprofit organizations offers a concrete logic model for performance measurement. However, the process is quite cumbersome and resource intensive to employ, complete, and integrate into normal operations. The model is a better fit to self-assessment, to identify key strengths and opportunities for process improvement and strategic change, and is more effective when supported by a core value of organizational learning (Ford & Evans, 2002). The Baldrige Award Criteria for Performance Excellence scores can be used comparatively or for benchmarking purposes, though the number of nonprofit organizations using the tool is unknown.

***Social return on investment.*** As part of an effort to make performance measurement more nonprofit-centric, strategies such as social return on investment (SROI) have been employed (S. R. Smith, 2010). With focus shifting from a charity mentality to one that involves philanthropy as a social investment, the demand to understand, communicate, and quantify the value creation of such giving led to the SROI (Emerson et al., 2000). Many evaluation methods omit the social value a nonprofit organization has within the community. The SROI “is designed to overcome this problem through a more inclusive approach to thinking about costs and benefits that consider the savings to society of nonprofit services” (S. R. Smith, 2010, p. 136). Value is created in economic, socioeconomic, and social continua. Social value is created when the lives of community members or society is improved (Emerson et al., 2000). However, most implementations of SROI maintain a programmatic level of impact analysis, rather than a holistic view of organizational impact within a community. The model is complicated in practice, and few nonprofit organizations have adopted the practice.

***Outcome measurement (internal evaluation).*** Performance measurement is a process of measuring inputs, outputs, and outcomes and the term is often used interchangeably with outcome measurement (Benjamin & Misra, 2006). As predominantly practiced within the nonprofit sector, performance measurement should shift in focus from processes (efficiency) and outputs (quantity served) toward specific outcome measurement (Lowell et al., 2001). An outcome “describes a specific desirable result or quality of an organization’s services” (Morley, Vinson, & Hatry, 2001, p. 5). Outcome measurement includes identifying desired outcomes, developing data collection

procedures and indicators, conducting data analysis, and reporting the findings (Morley et al., 2001).

Within the sector, leaders of very few organizations practice outcome measurement (Berger, 2009; Sawhill & Williamson, 2001). Nonprofit organizations are accustomed to providing output information, such as the number of clients served, but few track or understand how they help their clients. Program evaluation tools offer systematic methods for assessing results, strategic planning, and decision making (Kluger, n.d.) through maintaining the focus on the program rather than organizational level and further institutionalizing the practice of program-level evaluation and funding. The focus of most outcome measurement research and practice is on program output rather than organizational outcomes (Bell-Rose, 2004).

Foundations and other funders are increasingly encouraged to create social value through funding organizations that measure performance outcomes (Porter & Kramer, 1999) and in turn are making the same demands of fund recipients, which is explained by resource and institutional theories. Those highly dependent upon external funding sources will shift practices and begin outcome measurement to ensure continued funding (Thomson, 2010), which is evidence of isomorphism.

Philanthropy by private foundations was historically a form of gift giving, and performance measurement was not a priority; however, the shift in the definition of philanthropy toward that of social investment has increased demand for social return, impact, effectiveness, and sustainability (Leat, 2006). In many instances when performance measurement is required, funders and recipient organizations rarely discuss desired outcomes, benchmarks, or evaluation methods before the program is under way.

Without the intentional development of a clear understanding between both parties of performance expectations, how evaluation will occur, what defines strong or weak performance, and how results will be used, performance measurement will remain an elusive afterthought (Carson, 2000).

***Empirical research.*** Originally used and studied in the public sector, a growing body of empirical research is beginning to define and interpret performance measurement among nonprofit organizations. The studies help to define the extent of performance measurement use, current performance measurement practices, perceptions of performance measurement, and the challenges faced by nonprofit organizations implementing performance measurement (Thomson, 2010).

In a study of 101 nonprofit organizations in Franklin County, Ohio, 49% had reliable mission measures and only 14% used performance measurement (Sheehan, 1996). A study of 31 national nonprofit organizations found the same dynamic, with the majority focusing on output rather than mission-directed outcomes (Sawhill & Williamson, 2001). The increased interest in outcome evaluation can be attributed to the belief that outcome measurement will identify successful programs and, for grant makers and other funders, which programs should receive continued support (Carson, 2000). Even when nonprofit organizations have conducted outcome measurement and shown a program to be successful, funding can be fleeting as funders shift funding priorities or search for newer, more innovative programs rather than fund those with proven results (Carson, 2000). Fine, Thayer, and Coghlan (2000) studied 178 nonprofit organizations known for their evaluation practices and found 44% had conducted a program evaluation that included outcome measurement within 3 years. The most common reasons nonprofit

leaders conduct performance measurement include accreditation, financial audits, personnel evaluation; record reviews for regulatory bodies; and report to external parties (e.g., funders; Carman, 2007).

Program outcomes were evaluated by 62% of the 91 Dallas nonprofit organizations surveyed (Hoefler, 2000). Poole, Nelson, Carnahan, Chepenik, and Tubiak (2000) examined 191 performance measurement systems developed by 78 central Florida nonprofit organizations receiving United Way funding. Poole et al. found that although the majority had well-developed output measures in their performance measurement plans, only 60% identified outcome measures and included formal evaluation plans, and it is unknown how many of these plans were successfully implemented and used by the nonprofit organizations. A study of 149 nonprofit organizations in South Carolina indicated 65% used performance measurement in evaluation practices (Zimmermann & Stevens, 2006).

In a survey of 178 New York nonprofit leaders and 31 interviews, Carman (2007, 2009) found that leaders of organizations receiving funding from federal sources and United Way affiliates were more likely to conduct evaluation and performance measurement than were leaders of organizations not receiving this funding. Of those sampled, 65% conducted formal program assessments and 41% had performance measurement systems, 17% designed program logic models, and 3% used balanced scorecard management systems. Surveys revealed 83% of executives or management staff were responsible for evaluations, and fewer than 5% used external evaluators. More than 62% used internal operating funds to pay for evaluations, whereas 25% cited no evaluation expenses and in these cases interviewees revealed the leaders conducted

evaluation through financial measures and meeting budgetary projections. The respondents indicated 94% measured outputs and 71% measured service quality, and while most collected input and output data, few collected outcomes, customer satisfaction, or comparative data. The results reiterated the normative affect that funder expectations have on nonprofit practices and how measurement is predominantly defined at a program level. A study of 189 Indiana nonprofit organizations found performance measurement to be less common, with only 62% measuring program outcomes (Carman & Fredericks, 2010). Location might affect the level of implementation and sophistication of performance measurement practices. Morley et al. (2001) interviewed 36 nonprofit organizations and found that while 77% conducted measurement on some level, 44% asked only about client satisfaction, 44% used outcome measurement to improve services, and 28% used outcome measurement for fundraising purposes.

The majority of these studies did not assess the quality of the performance measurement measures. The difference between nonprofit organizational intentions and actual practice is unknown, and the use of output measurement was the predominant finding (Thomson, 2010). Thomson (2010) found, of the 237 Detroit nonprofit organizations receiving Community Development Block Grant funding in 2002, that 82% had at least two outcomes and 39% had four or more. Of these organizations, 8% expected to achieve long-term (end) outcomes, 38% expected short-term results, and 50% expected intermediate outcomes, yet only 55% used methods that would enable them to determine whether these outcomes were met. Forty-one percent did not have valid assessment tools for their identified outcomes and only 26% measured and reported outcomes regularly. The implication is that nonprofit leaders understand the importance

of outcome measurement, are complying with the demand to identify outcomes, and indicate their organization practiced outcome measurement, though few have the tools, resources, or capacity to measure and report outcome information appropriately.

***Limitations.*** Performance measurement focused solely on outcomes is flawed as the emphasis is solely on short-term improvements and neglects future performance because drivers are not identified or tracked (Holton, 1999). It is also a failure to assume that outcomes will be manifested during the short term in which many funders expect results. Many performance outcomes can take years to be delivered (Holton, 1999). Another limitation of outcome measurement is that organizations or programs are not necessarily the sole factor responsible for observed changes. Many other societal factors can contribute to the outcomes measured by nonprofit organizations (Henderson, Chase, & Woodson, 2002). Herman (1992) cautioned use of outcome indicators and multidimensional outcome measurement as sole determinants of organizational effectiveness because no agreed upon set of outcomes exists. Herman defined organizational effectiveness as “a matter of effective at what, for whom, and according to whom” (p. 415).

Outcome measurement requires clear, specific, and clarified goals and objectives; measures linked to the established goals; valid and reliable data collection, and time allotted for collection and analysis (Buckmaster, 1999). Outcome measurement has many inherent challenges for nonprofit organizations, including (a) the increased expense, (b) social outcomes are elusive, (c) and there is resistance within the NPO sector due to the punitive connotations of performance measurement (Easterling, 2000). Outcome measurement is limited by focusing solely on what has occurred and fails to

consider opportunity costs or indirect or unintended effects (Leat, 2006). Unless specifically designed into the system to counteract this, outcome measurement is historically focused rather than future oriented. To be successful, performance measurement must be linked to leadership and values, the strategic vision, stakeholders, and organizational culture; in addition, the interaction between internal and external relationships and factors must be monitored (Davenport & Gardiner, 2007). Outcome measurement is only one facet of successful performance measurement.

*Capacity building efforts.* Restricted funding and small size lead to inadequate infrastructure (Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004a). “To deal with the inadequate funding for administration, organizations resort to the strategies of low pay, make do, and do without that diminish organizational effectiveness” (Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy, 2004a, p. 3). These practices continue to compromise the organization’s capacity to achieve its mission, achieve long-term benefit to society, and limit capacity for performance measurement. Increased attention has been devoted to performance management, outcome measurement, and in turn necessary capacity building to make such efforts possible.

Successful performance measurement requires short-term resource allocation at a minimum and long-term investment and development to ensure integrated practice (Bell-Rose, 2004). In response to the varied ability of nonprofit organizations to conduct and report performance measurement findings, many community organizations, consultants, and donor foundations have invested in nonprofit organizational capacity (Bell-Rose, 2004). Capacity building is the ability of a nonprofit organization to fulfill its mission



effectively (De Vita & Fleming, 2001). The goal is to create self-sufficient organizations and communities. Funding or expertise has been devoted to increasing capacity of governance, infrastructure, self-assessment capabilities, technical and professional assistance, strategic and financial management practices, and so forth. Few capacity-building efforts and infrastructure supports have affected the content of performance management, as the focus has remained largely on program outcomes rather than shifting toward organizational performance measurement (S. R. Smith, 2010).

**Mixed evaluation (insourcing).** In response to the demand for performance measurement, the internal strain it can have on an organization, and the efforts to build nonprofit capacity, Miller, Kobayashi, and Noble (2006) proposed the mixed evaluation method of insourcing. With limited financial and personnel resources, turnover, insufficient training, and the need to focus on the mission, the leaders of many NPOs are limited in their ability to measure performance outcomes effectively or find performance measurement to be unsustainable. “Insourcing by our coin is a distinct mix of in-house and outsourced evaluation permitting sustained outcome assessment at a relatively low cost” (Miller, Kobayashi, & Noble, 2006, p. 86). Insourcing is a partnership between a nonprofit organization’s program staff and professional evaluators, where program staff members provide information and support while evaluators create the structure and analysis necessary to provide insightful data. The benefit of such a method is that organizational leaders can gain performance outcome data without sacrificing the mission or need to build capacity and organizational knowledge to do so. When similar types of organizations use this method, logic models and protocols can be developed for like-organizations, thereby reducing the cost for the method.

Unlike other methods, insourcing is a minimalist approach that does not require organizational staff to acquire evaluation skills (Alaimo, 2008). As such, insourcing is a compromise so evaluation can occur less expensively than building organizational capacity for self-evaluation and creates evaluation sustainability (Alaimo, 2008). A weakness of insourcing is the potential for funders to minimize or revoke funding for evaluation measures. The method is better suited for smaller to mid-size organizations.

**High-impact and social value measurement.** Organizational effectiveness has become increasingly important as a means for demonstrating social impact (Sowa et al., 2004). “Social value can be defined as the total social impact a charitable organization has on all its stakeholders and thus needs to have a broader scope than might traditionally be considered in performance evaluation” (Polonsky & Grau, 2008, p. 130). The start to defining high levels of impact is developing a link between organizational results and the inherent goals in an organization’s mission statement (Kelly & Lewis, 2009).

Using an inductive case-based methodology, Crutchfield and Grant (2008) sought to define the practices of high impact organizations and identified American nonprofit organizations formed between 1965 and 1994 that had served a broad public interest and achieved substantial and sustainable results at a national or international level in their short tenure. The four-part study identified 12 high-impact organizations and the six common practices employed within the organizations. The high-impact framework includes government advocacy, private sector collaboration, engaging and inspiring volunteers and donors, adaptation, building nonprofit networks and alliances, and shared leadership and empowerment.

Although the high-impact organizations achieved notable results, many did not have strong management systems, and despite developing strong brand awareness, many did not focus on marketing practices (Crutchfield & Grant, 2008). A compelling mission statement guided each organization, and each focused on impact and making a difference rather than inputs and outputs. Many scored low on financial efficiency standards. The high-impact framework provides a prescriptive set of practices, and while these practices served the 12 high-impact organizations, there is no guarantee that the same practices would lead other organizations in different contexts to the same level of results.

Current performance measurement and evaluation practices span a wide range of techniques. Performance and performance expectations are socially constructed concepts. There are many models, systems, and even subsystems that examine different dimensions and subsections of performance. Each of these methods has inherent strengths, yet they all fail to take a holistic view of nonprofit organizational performance, and few could address performance across organizational demographics and missions.

**Lack of consensus and disconnected momentum.** Many concepts exist for how to define and measure performance and effectiveness. The issue about whom nonprofit organizations should be accountable to, for what reasons, and how they should be measured appears in many conceptual frameworks, with few in agreement (Benjamin, 2008). The centrality of organizational effectiveness is generally agreed upon; however, there is little agreement on the definition, measurement, and constitution of effectiveness (Lecy et al., n.d.). Each new construction of nonprofit organizational performance affects how to provide resources and how to contrive expectations, which leaves the nonprofit

sector and its leaders in a continual tug-of-war between whose definition and method to implement and follow.

Many ideas exist for how to evaluate, but relatively few if any large-scale empirical studies include an examination of the evaluation practices of nonprofit organizations, and most studies are narrow in focus or are case studies examining the practices at a few organizations (Carman, 2009). More theoretical and conceptual frameworks have been presented than empirically tested. With the rich field of opinions and methods, it is not surprising that the empirical findings demonstrate the difficulty nonprofit organizations have in designing and implementing effective performance measurement systems, for what purpose, and collecting what information.

### **What Is the Purpose of Performance Evaluation?**

Ebrahim (2009) identified three streams of normative logic in nonprofit accountability: standard normative logic, performance-based logic, and mission-based logic. In the standard normative logic, accountability exists to increase oversight through regulation, disclosure, monitoring, and enforcement. The focus is on organizational governance and is characterized as coercive or punitive. In performance-based reporting, the logic assumes organizations should be accountable for delivering results through performance measurement. The focus is on analysis and progress toward goal achievement, is characterized as professional or technocratic, and is often self-regulatory though reported. The emergent mission-based accountability involves examining progress toward mission achievement. The focus is on systematic reflection and is characterized as strategic and adaptive. The three accountability logics are not mutually exclusive, and no formal hierarchy exists between the logics. “Framings of

accountability problems and their solutions tend to be driven by normative agendas rather than empirical realities” (Ebrahim, 2009, p. 891). Typically, accountability is framed in principal–agent terms and rarely includes joint problem solving. In its current conception, the purpose, content, and method of performance measurement is highly subjective and elusive.

### **Creating Effective Performance and Performance Measurement**

Although it is generally understood that performance information should be included in nonprofit management and assessment, limited adoption and use have occurred. This is a result of a shift from outputs to outcomes in performance measurement, performance measurement requiring fundamental management changes, resource demands, and many years and consistent effort to ensure change (Mayne, 2007). Challenges can be behavioral (political/cultural) or technical (measurement and reporting; Mayne, 2007; Thomson, 2010). For the shift to be successful, many transformational and change management leadership practices are advised. The practices include setting the climate and expectations, data selection, accountability, and avoiding distorting behavior (Mayne, 2007). To address technical challenges, measurement attribution systems must be developed, data must be linked between financial and performance information, and data quality and reporting credibility must be improved (Mayne, 2007).

**Internal characteristics.** Performance measurement, and the identification of outcomes, allows nonprofit leaders to understand how the organization is fairing, both good and bad, and allows for identifying what is working and what needs improvement (Morley et al., 2001). Sawhill and Williamson (2001) offered four lessons for effective

performance measurement: (a) measurable goals are required for mission impact measurement, (b) measures must be simple and easily communicated, (c) measures have inherent marketing value and stakeholders appreciate organizations that hold themselves accountable, and (d) measures should be used to manage the organization not just for assessment purposes. Performance measurement must be an integrated system for it to be useful, repeated, and of continued value for an organization.

***Systems thinking.*** The process of integrating performance measurement into management practices and decision making requires a change in culture and organizational learning (Mayne, 2007). An organization must have systems thinking to identify how interrelated practices and beliefs enable or prevent performance improvement and to create structures for employee empowerment, team learning, and a shared vision (Senge, 1990). An essential part of the process is the feedback loop. “Organizational feedback is defined as the amount of information related to organizational functioning received by an organization from an extra organizational unit(s)” (Morrow, 1982, p. 438). How the organization promotes, facilitates, and incorporates feedback is of critical importance to organizational learning and performance improvements.

***Feedback mechanisms.*** Effective performance measurement also incorporates single-, double-, and triple-loop learning. Single-loop learning occurs when a single feedback loop informs employees when there is a difference between expected and obtained outcomes (Argyris, 1976; Spitzer, 2007). The reflective practice focuses on how specific actions affect consequences. Double-loop learning occurs when employees question the content, beliefs, assumptions, and values that affect performance (Argyris,

1976; Spitzer, 2007). In double-loop learning, the practice becomes reflexive and focuses on how assumptions affect actions, in turn affecting consequences.

Organizations and employees become critically reflexive as a function of triple-loop learning. In practice, employees question context and how reality, knowledge, and understanding are socially constructed (Cunliffe, 2004). The practice allows for identifying organizational realities, contradictions, doubts, and dilemmas and can further inform the performance measurement development, implementation, and integration processes.

*Adaptive capacity.* Organizations with strong adaptive capacity function as learning organizations and use assessment as a means to increase organizational learning and enhance performance (Letts, Ryan, & Grossman, 1999; Strichman, Bickel, & Marshood, 2008; Sussman, 2003). Characteristics of adaptive capacity include shared vision, inquisitiveness and openness, evaluative thinking, systems thinking, social capital, external focus, and network connectedness (Strichman et al., 2008). Evaluation and assessment are participatory processes, and in these organizations, staff members are empowered and ample opportunities for learning and development are present (Strichman et al., 2008). Continuous learning, or the increased understanding and improvement achieved from performance measurement and assessment, allows for increased organizational effectiveness (Bailey, 2005).

*Appreciative inquiry.* A practice organizational leaders can use to engage employees to renew, change, or enhance performance is appreciative inquiry (Cooperrider & Whitney, 2005). Appreciative inquiry enables participants to focus on the connections, practices, and beliefs that enable success through focused questions,

collaboration, and systems thinking. Where evaluation can take negative connotations, appreciative inquiry focuses on creating positive organizational change.

**External characteristics.** Effective performance measurement, from an external perspective, is likely to be defined differently dependent upon the responding party. Simply stated, there are likely to be common if not generic characteristics of an effective performance measurement system. First, performance measurement would present information desired by upstream customers: donors, regulators, media, interest groups, and government parties (Moore, 2000). Second, data would be fact-based, relevant, readily available, and comparative in some fashion (O’Flanagan et al., 2008). Data would also be outcome focused, including both quantitative and qualitative information (O’Flanagan et al., 2008). Except in cases of legal or ethical wrongdoing, performance measurement should not be used for punitive purposes, but rather as a tool to increase dialogue between NPO leaders and upstream customers to provide the support necessary to improve performance and social impact.

**The leader’s importance to performance measurement.** One of the most underappreciated leadership roles is measurement leadership (Spitzer, 2007). If nonprofit leaders fail to make performance measurement a specific and intentional practice, “measurement in organizations will continue to become even more *dis-integrated, sub-optimizing, and non-transformational*. At the root of the problem are two misconceptions: 1) that measurement is primarily a technical activity, and 2) that no one individual has *overall* responsibility” (Spitzer, 2007, p. 121). Nonprofit leaders must take an active role in creating a culture of performance measurement and performance improvements if they are to advance the mission of the organizations they lead.



Performance measurement is one of the most critical practices of an organization to determine and expose value creation and destruction, and as such, organizational leaders must make measurement leadership a fundamental part of their jobs (Spitzer, 2007).

### **Perceptions of Organizational Effectiveness and Performance Measurement**

Cultural challenges within the nonprofit sector exist, as performance measurement practices adopted from the for-profit sector are typically seen as suspect. Nonprofit leaders “face the serious problem of how to balance tools from other sectors that may allow them to improve organizational performance within an organizational culture that is strongly suspicious of both private and public sector techniques and motivations” (Lindenberg, 2001, p. 248). Relevant stakeholders often perceive evaluation as unpleasant, threatening, and a forced requirement of outsiders (Donaldson, 2001, p. 356). Concerns include the expense of evaluation, the credibility of evaluators, and the negative consequences resulting from evaluation (Donaldson, 2001). The internal and external emphasis on the nonprofit sector “doing ‘good work’ creates an environment where the struggles, failures, or mistakes seem less likely to be a part of the conversation” (Benjamin & Misra, 2006, p. 156). An inability to discuss failures and areas for improvement prevent organizational learning and maintain cultural resistance toward performance measurement.

Cultural resistance toward performance measurement is complicated by the varying performance measurement definitions and practices imposed upon nonprofit leaders that are ill fitting or ill conceived. Forbes (1998) commented, “The inherent subjectivity of effectiveness assessment has probably always been more apparent to people who work with NPOs on a daily basis” (p. 197). The negotiable and malleable

construction of organizational effectiveness creates incoherent direction for nonprofit leaders, threatens plans and strategy, and threatens service stability (Forbes, 1998). Most performance measurement is externally driven and as a result often fails to include nonprofit leaders in the development of standards, lacks critical buy-in, reports information that is not representative of actual results or practices, and rarely assists nonprofit leaders in making performance improvements.

**Current related research.** Studies on nonprofit leader perceptions of performance measurement are rare. Most involved collecting perceptions while studying performance measurement practices or focusing on a topic related to performance measurement. Such studies include those that examine performance in relation to the individual leader's characteristics (Ritchie, Anthony, & Rubens, 2004), nonprofit leader perceptions regarding funding (Huang, Buchanan, & Buteau, 2006), perceptions of program evaluation capacity (Alaimo, 2008), or captured leader perceptions while studying evaluation practices (Carman, 2007; Sawhill & Williamson, 2001).

***Performance in relation to leader characteristics.*** To determine the extent that individual leader differences create the variance in executive perceptions of performance measurement, Ritchie et al. (2004) found collectivism, preference for analysis, and locus of control explained the variance in performance measurement perceptions held by 144 foundation chief executives. The results indicated the stance an organization takes on performance measurement relates to the personality and preference of its leader. It is unknown whether this finding is replicated in nonprofit organizations or how this influences the purpose, content, or process elements of nonprofit performance measurement.

***NPO leader perceptions of funding.*** In a study for the Center of Effective Philanthropy, Huang et al. (2006) examined foundation and grantee perspectives regarding foundation support and the impact of program versus operational support. The study included nearly 15,000 grantee-completed surveys rating 163 large foundations, surveys from the CEOs of the 163 foundations, and interviews with 26 leaders from grantee organizations. The results showed most grants are program restricted, small, and short term, which is counterintuitive to most foundation CEOs believing that operating support is more likely to have a positive effect for grantees. Fewer than 16% of foundation CEOs favored operating support, and of the 49% who preferred programmatic support, 30% cited the ease in measuring outcomes as the reason for their preferred support. Grantees preferred operating grants when amounts were larger and had longer terms than typically awarded, yet the median operating grant was \$50,000 with a 1-year term. Foundation CEOs found operating grants more difficult to assess and cited reasons for heavy program support including board pressure, fit with the foundation's mission, unfamiliarity with the grantee, and concerns over creating grantee dependency. "There is agreement among CEOs on at least one thing: operating support is viewed as being most effective—and more effective than program support—in creating impact and on encouraging sustainability of grantee organizations" (Huang et al., 2006, p. 10).

Huang et al. (2006) revealed that tension exists between what is best for the foundation and what is best for the grantee. Resolution more often sides in favor of program support, as it better meets foundation needs than the needs of the community. According to the foundation CEOs, this preference for program support occurred mostly because NPOs could not quantify or assess organizational performance. Acting in the

name of assistance, outside funders often make operations more difficult for nonprofit leaders, failing to provide the assistance that is most beneficial in favor of their own needs.

***Perceptions of program evaluation capacity.*** Alaimo (2008) conducted 126 interviews with nonprofit leaders, board chairs, program staff, funders, and evaluators from Atlanta and Indianapolis to ascertain their perspectives regarding the program evaluation capacity of 42 participating health and safety nonprofit organizations. Participants were grouped into three types. Type I organizations (31%) used comprehensive evaluation methods including program-driven and outcome-focused methods conducted by dedicated staff or external evaluators and demonstrated a long-term commitment to evaluation practices and process improvement. Type II (67%) organizations employed a single evaluation method to comply with customer satisfaction and external requirements, conducted ad hoc by nondedicated employees, and focused on process outputs with sporadic commitment to evaluation practices and process improvements. Type III (2%) organizations exhibited no effort to evaluate programs.

Alaimo (2008) discovered that Type I organizations had a median age that was 40 years older and median revenues 12 times higher than Type II and Type III organizations. The older the organization and the higher level of resources, the more likely that the program evaluation was institutionalized, outcome driven, and improved to deliver pertinent data. Nonprofit leaders at Type I organizations achieved higher levels of education (69% graduate, 31% undergraduate) than their Type II (46% and 46%, respectively) and Type III counterparts (100% diploma), and 100% received evaluation instruction versus 68% of Type II and 0% of Type III. Education levels appear to play a

pivotal role in the development and integration of performance measurement practices at these organizations.

According to Alaimo (2008), “Of the 48% of the EDs who chose program effectiveness as a top three priority, 33% ranked it as their first priority, 54% as their second priority and 13% as their third priority” (p. 210) and Type I organizations listed program effectiveness within their top two priorities. Evaluation was funded from operating funds at 88% of the organizations (Alaimo, 2008). Conversely, 30% of board members listed financial stability as their top priority, 15% listed mission achievement, and 38% of board chairs ensured program evaluation was included in strategic plans (Alaimo, 2008). External evaluators overwhelmingly revealed the primary challenge in successful program evaluation was organizational culture.

Although 100% of funders in Alaimo’s (2008) study requested information about program activities (outputs) and 93% required financial information (efficiency), only 67% required program evaluation information (outcomes). Of the 67% requiring program evaluation, 100% requested outcome data, 38% regularly provided financial support for program evaluation, and 20% believed their counterparts funded program evaluation (Alaimo, 2008). Funders who request outcome data do not want to fund such evaluation practices, and few believe that other funders do either. With funders limiting operating funding, the number of organizations conducting performance measurement and the quality of those with performance measurement systems is not surprising. Without adequate funding, nonprofit leaders will continue to be constrained by poorly designed and implemented performance measurement systems generating suspect, low-quality data.

*Leader perceptions captured while studying evaluation practices.* Sawhill and Williamson (2001) proposed impact, activity, and capacity for performance measurement, which aligns with organizational mission, vision, goals, strategies, and programs. Interviews with 30 CEOs from national and well-known nonprofit organizations showed the repeated response of difficulty in effectively measuring mission success (Sawhill & Williamson, 2001). Most of the organizations had the nation's largest operating budgets and available resources yet still have difficulty developing effective performance measurement systems. As this is a problem for organizations with the highest levels of resources, it provides insight into why this problem is rampant for the majority of the sector that is often more resource dependent. Coupled with the wide range of expectations from various stakeholders, it is not surprising that many organizational leaders do not know how or what to measure. The leaders of most of these mature organizations had methods for measuring the activity and capacity, as these areas were more likely to include measurable goals. The more abstract the mission, the more difficult it is to measure (Sawhill & Williamson, 2001) and the less likely the organizational leaders understand whether it is making an impact.

In 31 interviews and surveys of 300 New York nonprofit organizations, Carman (2007, 2009) uncovered that while 74% developed performance targets, how these were developed varied widely, with some interviewees admitting their targets were “‘fiction’ or ‘made up’” (Carman, 2007, p. 66). Others revealed *outcomes* was the new, more fashionable word for goals, and many were struggling with what they called the “‘outcome movement’” (Carman, 2007, p. 66). Carman (2007) observed common themes in the difficulty to measure outcomes, external pressure to measure, and unless

performance measurement was institutionalized or without technical assistance, the practice was illusive and cumbersome. Nearly every organizational leader viewed quality control and accountability as the purpose for performance measurement. Carman (2007) revealed that nonprofit leaders provide evaluation data to suit the requests of external parties. To improve nonprofit organizational evaluation practices, Carman (2007, 2009) suggested that external parties change what is expected and requested, lend technical assistance, and develop evaluation capacity.

*Performance measurement characteristics.* In a rare case study about performance measurement purpose, content, and process, Moxham (2009) uncovered perceptions of performance measurement from managers of six United Kingdom-based nonprofit organizations and their associated funders and regulators. The leaders identified the purpose of performance measurement was for external accountability, to facilitate financial reporting, to demonstrate achievements, for operational control, and a few revealed to facilitate continuous improvement. The leaders noted government policy, funders, regulators, and the nonprofit organizations developed performance measurement criteria out of frustration for external measures not fitting the organizational efforts. The nonprofit leaders noted that performance measurement occurred through self-assessment, auditors, and informally. When asked to define outcomes, few could do so even after reading a brief description of outcome measurement. The existence of the range of performance measurement perceptions is not surprising because both internal and external parties were included in the study. Although Moxham uncovered performance measurement perceptions, the study was limited in size, location, and applicability to the nonprofit sector.

***Performance measurement practices.*** Hetrick (2004) studied 230 member nonprofit organizations from seven medium-sized United Way of America regions to identify performance measurement practices, in addition to nonprofit leaders' perceptions of how their organizations fared at performance measurement. Despite 90% of the nonprofit leaders possessing an undergraduate degree, only 33% considered themselves highly knowledgeable about outcome measurement and performance measurement. Of the respondents, 45% indicated the organization received government funding for programs, though nearly 32% viewed United Way of America as a primary source for assistance with outcome measurement and performance measurement, which outranked a national or parent organization for this responsibility. The most common source of outcome measurement training was offsite workshops (82%); however, only 23% found the training to be very helpful, 44% thought the training was somewhat helpful, and 10% found the training to be not much help. The assistance ranked highest in helpfulness was funds for training (40%).

Further, in terms of annual performance measurement practices, Hetrick (2004) found 45% reviewed their mission statements, 76% reviewed goals and objectives, 89% evaluated program outputs, and 84% evaluated program outcomes. Despite only 33% considering themselves highly knowledgeable about outcome measurement and performance measurement, the leaders of a surprising and overwhelming majority of organizations were confident in their outcome measurement practices. Although 84% noted they evaluated program outcomes, when asked about capturing long-term outcome data, only 15% thought they did this very well, 36% responded somewhat well, 34% selected not too well or not well at all, and 14.5% did not measure this factor at all. The



leaders were also asked why employees do not conduct performance measurement, and 44% stated that employees were not expected to measure outcomes, 26% cited lack of time as a reason, and 27% listed lack of training.

In terms of data analysis, Hetrick (2004) found 89% of the nonprofit leaders said results were compared to the previous year, 82% compared results to established targets, 47% evaluated results against accepted standards for the field, and 18% compared their results to high-performance nonprofit organizations. When a program does not meet performance goals, 42% suggested increasing human resources assigned to the program, 23% phase the program out, 16% increase funding, 16% take no action, and 15% decrease funding. The quality of performance measurement data is suspect when 58% throw more resources into a flawed program, 23% eliminate the flawed program, and 16% do nothing. Despite the overwhelming majority reporting outcome measurement practices at their organization, nearly half later thought their organization was inadequate at collecting outcome data or did not collect the data after all.

**An argument for leader perceptions of performance measurement.** A sparse number of studies examining nonprofit leader perceptions of performance measurement factors exist. This study involved identifying the purpose, content, and process factors that nonprofit leaders believe to be the most important for performance measurement, which is different from current studies that examine how performance measurement happened, not how it should be conducted. An empirical understanding of how nonprofit leaders perceive performance measurement will help to identify whether commonalities in perception of performance measurement exist, how nonprofit leaders think performance measurement should be conducted, for what purposes, and for which

reasons. Challenges and roadblocks to successful implementation and integration of performance measurement could be reduced and widespread adoption could be facilitated after developing a common understanding of the performance measurement purpose, content, and process.

Individuals have deep-rooted mental models composed of internal images of how the world works (Senge, 1990; Starkey et al., 2004). These mental models can have a powerful influence on individual action because of their effect on how an individual perceives the world. Understanding nonprofit leader perceptions of performance measurement might help to uncover challenges in achieving a wide-scale replacement for financial efficiency in the evaluation of nonprofit organizations. Individual perception affects individual learning and decision-making processes, which is the foundation of organizational action and learning (Starkey et al., 2004). Understanding nonprofit leader perceptions might help to uncover how nonprofit organizations are positioned for organizational learning and making performance improvements.

### **Review of Methodologies**

Most nonprofit organizational performance measurement research has been theoretical in nature. Empirical nonprofit organizational performance measurement research includes qualitative case studies and interviews, quantitative longitudinal studies and surveys, and mixed-method designs incorporating quantitative surveys with qualitative case studies or interviews. The selection of research design includes a set of decisions made by the researcher about the topic under study, among which are population, research method employed and for what purpose (Babbie, 2008), and

providing the plan for data collection, measurement, and analysis (Cooper & Schindler, 2008).

In nonexperimental research, research designs can be causal, descriptive, and exploratory (Salkind, 2008). Descriptive methods involve determining the specific characteristics of a phenomenon existing within a population, providing a broad depiction of the research topic (Salkind, 2008). Causal or explanatory designs involve an attempt to explain why a phenomenon occurs and what the causes or reasons are for the occurrence of the topic of study (Neuman, 2006). Exploratory designs involve an attempt to address new topics or those with little previous research to “formulate more precise questions that future research can answer” (Neuman, 2006, p. 29).

An exploratory design is appropriate for this study as the goal of the research is to identify nonprofit leader perceptions regarding performance measurement, rather than to explain or establish causal relationships. A researcher can use many methodologies to study nonprofit leader perceptions of performance measurement; however, the research problem and research questions should indicate the method chosen (Creswell, 2005). Exploratory studies can include quantitative, qualitative, or mixed methodologies. Survey questionnaires, interviews, focus groups, and document analysis have been employed to study and articulate nonprofit performance measurement practices and techniques. The following review includes possible methods that were considered for the study.

**Quantitative research methods.** Quantitative research methods involve an attempt to measure observable facts, focus on variables, emphasize reliability, function independent of context, observe many cases or subjects, utilize statistical analysis, and

attempt to remove the researcher from the process (Neuman, 2006). Researchers of quantitative studies collect data in the form of numbers through experiments, surveys, content analysis, and existing statistics (Neuman, 2006). The most common quantitative research designs are experimental, correlational, and survey research. Correlational research methods identify how two or more variables are related, what they have in common, and how to predict specific outcomes with this information (Salkind, 2008). Researchers of experimental studies seek to identify the outcomes experienced by participants through specific interventions (Creswell, 2009). Survey research enables researchers to produce statistics about a targeted population, describing the target population's characteristics because of the design enabling probability sampling, standardized measurement, and customization (Fowler, 2009).

If this study were to be a quantitative study, a survey instrument would be most appropriate for collecting and analyzing nonprofit leaders' perceptions of performance measurement. Although survey procedures allow for describing and generalizing the characteristics of a large population, survey research rarely captures the context of social life, is subject to artificiality due to participant self-reporting, and cannot be adapted or changed throughout the study to address variable operating conditions of a phenomenon (Rubin & Babbie, 2010). Additionally, "standardized questionnaire items often represent the least common denominator in assessing people's attitudes, orientations, circumstances, and experiences" (Rubin & Babbie, 2010, p. 404); because questions are designed to be at least minimally appropriate to all participants, often the questions can miss what is most appropriate to many participants.

**Qualitative research methods.** Whereas quantitative studies involve addressing trends and following a large body of previous research, qualitative research is appropriate when data are not readily available (Creswell, 2005). The most common qualitative research designs include grounded theory, ethnographic, and narrative research designs. Qualitative research methods involve an attempt to construct social reality and cultural meaning and to focus on interactive processes and events, emphasize authenticity, are constrained by the situational context, observe few cases or subjects, utilize thematic analysis, and inherently involve the researcher in the process (Neuman, 2006). Qualitative research helps to explain a central phenomenon, which might be a new concept or idea that has not been the subject of much academic endeavor (Creswell, 2005). Qualitative research involves examining human behavior within the social, political, and cultural contexts in which it occurs (Salkind, 2008) in the form of words or pictures through field and historical-comparative research (Neuman, 2006).

In qualitative research, the researcher “relies on the views of participants; asks broad, general questions; collects data consisting largely of words (or text) from participants; describes and analyzes these words for themes; and conducts the inquiry in a subjective, biased manner” (Creswell, 2005, p. 56). As such, there is a greater chance for the researcher to embed bias and affect participant responses during research interactions, largely because qualitative methods include interviews, focus groups, and other interactional formats whereas quantitative research methods generally remove the researcher from participants through the use of alternative research tools. Although qualitative methods allow the researcher to ask new questions and uncover new themes,

qualitative population samples are small, and results are not generalizable to the broader population (Cooper & Schindler, 2008).

**Mixed-method research.** Mixed-method research affords a researcher the strengths of both quantitative and qualitative forms, and in addition to collecting and analyzing both types of data, the use of both approaches in tandem creates a stronger study than quantitative or qualitative research alone (Cohen, Manion, Morrison, & Morrison, 2007; Creswell, 2009). A much richer understanding of nonprofit leader perceptions of performance measurement will be derived from a mixed-method, exploratory research design as it allows for the integration of participant subjectivity with statistical analysis of the results.

Tashakkori and Teddlie (2010) offered nine specific characteristics of mixed-method research. Mixed-method research offers methodological eclecticism, integrating the most appropriate techniques to investigate a phenomenon. Characterized by paradigm pluralism, mixed-method research utilizes a variety of paradigms as the foundation of inquiry. In mixed-method research, the diversity in the research allows for divergent conclusions and inferences while emphasizing a continuum of thought rather than dichotomies. Mixed-method research offers an iterative, cyclical research procedure where the research question guides the methods employed, characterized by unique design and analytical processes. Mixed-method research tends to offer balance and compromise implicitly, and visual representations are a common way to present findings (Tashakkori & Teddlie, 2010).

Common mixed-method research designs include triangulation, explanatory research, and exploratory research. Where triangulation involves collecting quantitative

and qualitative data simultaneously and then merging the data to understand the problem, explanatory research involves collecting quantitative data first followed by qualitative to help the researcher understand the picture presented in the numbers (Creswell, 2009). Exploratory designs involve collecting qualitative data first to investigate the phenomenon and then collecting quantitative data to help the researcher describe the relationship found in the qualitative data (Creswell, 2009).

**Q methodology.** Q methodology incorporates the strengths of qualitative and quantitative research methods, and through the systematic examination of subjectivity, might reveal connections overlooked by other techniques (Akhtar-Danesh et al., 2008). This study helped to uncover the subjective perceptions nonprofit leaders hold for the purpose, content, and processes for performance measurement. Q methodology provides for systematic analysis of human subjectivity (McKeown & Thomas, 2013) and helps to uncover what individuals believe and think about the topic of study (Sickler et al., 2006). Nonprofit leaders can help clarify the vast amount of information related to nonprofit organizational performance measurement and identify the most important purpose, content, and process factors to yield valuable data for external parties and internal performance improvement alike. Subjectivity is fundamentally an individual's point of view, and in contrast to typical research that resonates from the external standpoint of the researcher, Q methodology allows for exploring participants' internal standpoint (S. R. Brown, 1980).

## **Conclusion**

As the nonprofit sector continues to grow in size and as competition for declining resources becomes fiercer, external parties demand increased accountability and proven

results from nonprofit organizations (Ebrahim, 2005). Without an established definition of organizational effectiveness or a method for performance measurement, a widening concern is that the media and charity watchdogs are encouraging “donors to make decisions based on simplistic measures that can never substitute for real assessments of whether a nonprofit group makes a difference” (Wing, 2009, p. 13).

The external demand for accountability and measurement is complicated by an extensive body of literature defining a vast number of purposes, methods, and processes for measuring performance and organizational effectiveness (Forbes, 1998; Herman & Renz, 2008; Judge, 1994). Nonprofit leaders are increasingly subjected to socially constructed definitions for measurement and behavior (Herman & Renz, 1997), and because of their resource dependence, nonprofit organizations often conform and institutionalize the current demands and needs of external parties rather than focus on methods to best accomplish their missions (Froelich, 1999). Under the guise of assistance, external parties often make social impact a more elusive effect to achieve.

As such, nonprofit leaders often state their organizations practice outcome measurement and performance measurement, though few demonstrate systems that are capable of performing such a task. Current practices often hinder organizational learning or actual performance improvement (Lowell et al., 2001; Pallotta, 2008), which leads to conclusions that many current performance measurement practices are not in the best interest of nonprofit organizations but rather are a by-product of the needs of self-interested external parties. Externally imposed performance measurement usually fails to gain buy-in from nonprofit leaders and essentially reiterates the internal cultural resistance to evaluation (Spitzer, 2007). Without understanding nonprofit leaders’



perceptions of the purpose, content, or process elements that will best facilitate performance measurement at their organizations, performance measurement practices will likely continue to be ill conceived or ill fitted for measuring performance and effectiveness at an organizational level.

### **Summary**

The problem addressed in this study is the lack of a clear and consistent definition of organizational effectiveness and the inability of most performance measurement systems to measure nonprofit organizational performance effectively. Waves of interest in organizational effectiveness and nonprofit organizational performance have led to a variety of theoretical definitions but few empirically tested and proven models (Lecy et al., n.d.). The current efficiency standards hinder effective nonprofit organizational performance as the true costs of operations are often hidden or unrealized and investments in infrastructure are avoided to avoid punitive results (Ebrahim, 2005; Pallotta, 2008; Ritchie & Kolodinsky, 2003). As a result of their resource dependency, nonprofit organizations conform to and institutionalize socially constructed and imposed performance and measurement expectations that further complicate mission achievement (Froelich, 1999). Each of the resulting definitions and methods to measure organizational effectiveness and performance measurement are limited in that emphasis is often one dimensional, and a holistic view of organizational effectiveness and performance measurement has not yet been identified. The overwhelming question that remains about performance measurement is who should hold nonprofit organizations accountable, for what reasons, and how to measure performance measurement (Benjamin & Misra, 2006). The purpose of the Q methodology study was to gain a better understanding from

nonprofit leaders concerning their subjective perceptions of effective performance measurement practices.

Chapter 3 contains a review of the research conducted in this Q methodology study. The chapter will include a discussion of the design for the study, the appropriateness of the design, the population and sampling method, and the rationale for the data collection process. The chapter includes an examination of research design validity, the informed consent process, instrumentation, and data analysis.

## Chapter 3

### Method

The nonprofit sector is different from the public and private sectors in that organizational goals are typically intangible and more difficult to assess. Internal and external performance measurement methods have been proposed and tested, yet few, if any, have been applicable across a broad range of organizational demographics and characteristics. Ample dialogue exists for how to evaluate nonprofit organizations, yet there is a lack of consensus on how to best measure so that performance evaluation is relevant to external parties and assists nonprofit leaders with performance improvement. Performance measurement as practiced by the nonprofit sector tends to have an external focus, is retrospective, and rarely informs changes in practice (Emerson et al., 2000).

The purpose of this Q methodology study was to gain a better understanding of the subjective perspectives of nonprofit leaders in the United States concerning effective organizational performance measurement. Identification of the social perspectives held by nonprofit leaders may be the starting point in creating a more widely accepted and utilized performance measurement model that could supplement or replace financial efficiency as the benchmark for nonprofit effectiveness. Understanding and identifying the characteristics that each social perspective considers most important in nonprofit performance measurement could facilitate the creation of a model that has broad application across the diverse nonprofit sector.

Q methodology provides for systematic analysis of human subjectivity about a topic of study (McKeown & Thomas, 2013). Q methodology was appropriate for the study because it offers a way to quantify subjective perceptions and to identify patterns of

beliefs among nonprofit leaders (S. R. Brown, 1980; Swetnam, 2010). “Q identifies the views of groups who share a similar perspective on an idea or a topic by producing several factors of opinions” (Choi, 2010, p. 129). The study involved three groups of performance measurement elements: purpose (why organizations are evaluated), process (how organizations are evaluated), and content (what is evaluated). A chain referral sample of nonprofit leaders helped clarify the vast amount of information related to nonprofit performance measurement; identify the most important purpose, content, and process elements; and uncover the social perspectives that exist about this topic.

The research questions guiding the Q methodology study were as follows:

RQ1: What do nonprofit leaders believe to be the most important elements of effective nonprofit organizational performance measurement?

RQ2: What are the characteristics, behaviors, or practices nonprofit leaders believe are critical to developing an effective nonprofit organizational performance measurement model?

Chapter 3 explores the plan and methodology for the current research, in addition to discussing the appropriateness of the study design and research population. An overview of Q methodology, the selected methodology for the study, is included. Also reviewed are the processes for informed consent, confidentiality, data collection, instrumentation, data reliability and validity, and data analysis.

### **Research Method**

A researcher can use many methodologies to study nonprofit leader perceptions of performance measurement; however, the research problem and research questions should dictate the method employed (Creswell, 2005). Most nonprofit organizational

performance measurement research has been theoretical in nature. Empirical nonprofit organizational performance measurement research has employed qualitative case studies and interviews, quantitative longitudinal studies and surveys, and mixed-method designs incorporating quantitative surveys with qualitative case studies or interviews. Although quantitative survey procedures allow for describing and generalizing the characteristics of a large population, survey research rarely captures the context of social life, is subject to artificiality due to participant self-reporting, and cannot be adapted or changed throughout the study to address variable operating conditions of a phenomenon (Rubin & Babbie, 2010). Though qualitative methods allow the researcher to ask new questions and uncover new themes, qualitative population samples are small and results are not generalizable to the broader population (Cooper & Schindler, 2008).

The selection of research design includes a set of decisions made by the researcher about the topic under study; the population in which to study the topic; the research methods employed; the purpose of the study (Babbie, 2008); and the plan for the data collection, measurement, and analysis (Cooper & Schindler, 2008). In nonexperimental research, research designs can be classified as causal, descriptive, and exploratory (Salkind, 2008). Exploratory was the appropriate approach for this study as the goal of the research was to identify nonprofit leaders' perceptions regarding performance measurement, rather than to explain or establish causal relationships. Exploratory studies can use quantitative, qualitative, or mixed methodologies, and the review in Chapter 2 led to the conclusion that Q methodology was the most suitable method for this study. Q methodology incorporates the strengths of qualitative and

quantitative research methods, and through the systematic examination of subjectivity may reveal connections overlooked by other techniques (Akhtar-Danesh et al., 2008).

Q methodological research affords the researcher the strengths of both quantitative and qualitative forms, and as more than collecting and analyzing both types of data, the use of both approaches in tandem creates a stronger study than quantitative or qualitative research alone (Creswell, 2009). A much richer understanding of nonprofit leader perceptions may be derived from an exploratory, Q method design that facilitates the integration of participant subjectivity with factor analysis of the results.

### **Appropriateness of Design**

This study helped uncover the subjective perceptions nonprofit leaders hold for effective nonprofit performance measurement. Q methodology provides for a systematic analysis of human subjectivity (McKeown & Thomas, 2013) and helps to uncover what individuals believe and think about a topic (Sickler et al., 2006). Subjectivity is fundamentally an individual's point of view and opposed to typical research that resonates from the external standpoint of the researcher, and Q methodology allows for the exploration of participants' internal standpoint (S. R. Brown, 1980). "Q methodology is our window to learning about subjective reactions or responses to the issues confronting professionals daily" (Montgomery, 2010, p. 1).

The intent of Q methodology is to uncover common social perspectives about the topic of study. Q methodology researchers collect "a small number of distinctly different individual perspectives and then employ nonparametric statistics to reveal the underlying social perspectives" (Sickler et al., 2006, p. 356). Nonprofit leaders helped to clarify the vast amount of information related to nonprofit organizational performance measurement,

and through identifying the most salient purpose, content, and process elements, they revealed common social perspectives of nonprofit organizational performance measurement. “Q methodology differs from surveys in that it looks at the respondents as the variables, while the statements to which they respond are equivalent to ‘subjects’ (virtually inverting the survey method, which sees each question as a variable)” (Sickler et al., 2006, p. 356).

Q methodology, by design, is not suitable for all types of research. Q methodology “has the rather obvious disadvantage that it is designed to study subjectivities and thus is not suited to study matters of fact” (de Graaf & van Exel, 2008, p. 70), which makes Q methodology the ideal methodology for exploring participants’ preferences, motives, tastes, feelings, and understanding of a topic, the underlying parts of personality that influence behavior (van Exel & de Graaf, 2005). An individual’s conception of the topic, how the topic is understood and perceived, is defined in the Q sort (S. R. Brown, 1980). “As such, it is subjective and self-referent. It is operant in that it is in no way dependent on constructed effects. There is no right or wrong way to do a Q sort” (S. R. Brown, 1980, p. 6).

Researchers have employed survey questionnaires, interviews, case studies, and longitudinal analysis to study and articulate nonprofit performance measurement practices and techniques. Although survey procedures allow for the characteristics of a large population to be described and generalized, survey research rarely captures the context of social life, is subject to artificiality because of participant self-reporting, and cannot be adapted or changed throughout the study to address variable operating conditions of a phenomenon (Rubin & Babbie, 2010). Interviews and focus groups allow

the researcher to ask new questions and uncover new themes; however, qualitative population samples are small, and results are not generalizable to the broader population (Cooper & Schindler, 2008).

Q methodology is an inductive yet systematic methodology that allows values and preferences held by a population to be uncovered and analyzed using statistical methods (Steelman & Maguire, 1999). Problems associated with questionnaires such as central tendency and leniency are avoided through the required forced quasi-normal distribution into which participants sort statements (Khosrowpour, 2002), which allows for clarification of points of agreement and difference between participants (Valaitis et al., 2007). In Q methodology, the effects of researcher bias typical in qualitative designs are minimized because the subjective perceptions of participants are collected through the Q sort process, which is conducted independent of the researcher. “Q avoids the measurement problems encountered by researchers utilizing a priori meanings and categorizations” (Hathcoat & Montgomery, 2010, p. 42) because factors are defined after participants complete the Q sort.

During a Q sort, under a forced-choice condition of instruction, participants rank order statements’ order of significance with respect to their point of view (S. R. Brown, 1980, pp. 5-6). While the sample of persons sorting is necessarily small, is nonrandom, and may not be broadly applicable to individuals with different experiences (Akhtar-Danesh et al., 2008), the goal of this Q methodology study was to identify different patterns of thought rather than the numerical distribution within a larger population (Akhtar-Danesh et al., 2009). Correlation, factor analysis, and factor scores were used to analyze participants’ Q sorts (McKeown & Thomas, 2013) to identify common



performance measurement characteristics that nonprofit leaders believe are most important to the evaluation of nonprofit performance. PQMethod 2.33 software facilitated analysis of Q sort results, computing the Q sort data (Schmolck, 2002). Because the software is more conducive to Q analysis than standard statistical software platforms such as SPSS, PQMethod is the preferred and predominantly used data analysis software in Q studies (Webler et al., 2009).

As explained by S. R. Brown (1980), Q methodology's operant subjectivity differs from operational definitions of scaling and questionnaire methodologies for three main reasons. First, "a subjective operant, unlike a scale response, is neither right or wrong" (S. R. Brown, 1980, p. 4), creating little need for external validity because there is no outside criterion for a person's point of view. Second, where scales or tests require constructed effects, operants do not. "Operational definitions place constraints on behavior by replacing the subject's meaning with the investigator's, and the investigator ends up studying the constraints rather than the behavior" (pp. 4-5). Hence, the researcher creates a taxonomy of tests rather than a taxonomy of behavior (Stephenson, 1953). Last, operant approaches proceed in the opposite order of the construction of operational definitions. In operant subjectivity, a phenomenon is first observed and then concepts are attached to it (S. R. Brown, 1980, p. 5). "Operational definitions begin with concepts in search of behavior; operant definitions begin with behavior in search of concepts" (S. R. Brown, 1980, p. 28).

Validity and reliability are defined differently in Q methodology because the expressed perspective is derived from the participant's viewpoint rather than the researcher's external standpoint (S. R. Brown, 1980; Hathcoat & Montgomery, 2010).

Because the unit of measurement is the importance of a statement relative to all other statements according to the subject, conventional reliability and validity are not central to Q (Osterkiil, 2010). Validity is only an issue when participants fail to represent their own perspective (Choi, 2010; Hathcoat & Montgomery, 2010). Reliability is comprised of two parts: the test–retest coefficients from individual sorts and the schematic reliability or the replication of perspectives across multiple studies (Hathcoat & Montgomery, 2010).

Another criticism of Q methodology questions the reliability and the subjectivity of the researcher in relation to the process of statements chosen from literature and sorted by participants (Thomas & Baas, 1992). Strategic sampling helps reduce researcher bias in the Q-set creation, as statements are methodically chosen to represent the concourse both theoretically and inductively (Webler et al., 2009). In addition, previous studies have concluded that a limited number of unique viewpoints exist for a given topic (S. R. Brown, 1980), allowing for study replication. Replication, regardless of population, is a critical component for proving reliability (Buchman, 2009). Other researchers have found Q sort test–retest reliability to be 0.80 or higher, and content validity is assessed through a literature review and a review by domain experts (S. R. Brown, 1980; Valaitis et al., 2007).

Q methodology marries quantitative factor analysis with the detailed qualitative inquiry of human subjectivity. “Methodologically, Q gives substance to the logic of preference by explicitly recognizing the central role of subjectivity involved in evaluations of all kinds” (S. R. Brown, 1980, pp. 53-54). As such, Q methodology allows the researcher to gain a deeper understanding of the feelings, choices, motivations, and values held by participants than other methodologies provide (Stephenson, 1953).

## **Q methodology**

Defining and collecting the concourse is the first step in Q methodology. The concourse is a collection of all possible statements relevant to the topic of study, which contains all related aspects of all discourse about the topic, and comprises the raw material for Q methodology (van Exel & de Graaf, 2005). The statements comprising the concourse are collected from interviews, text, and other artifacts related to the topic. “The sample of statements must represent all key aspects of all the relevant perspectives on the issue and are selected in such a way that researchers do not impose their attitudes on the study” (Sickler et al., 2006, p. 357). Depending on the topic of study, the number of statements in the concourse will vary; however, the statements should be comprehensive and encompass all aspects of the topic (Akhtar-Danesh et al., 2008).

The concourse for this study was collected from the vast pool of dialogue about nonprofit performance measurement that exists in primary and secondary sources, many of which were reviewed in Chapter 2. The sources included academic journals, professional and trade publications, books, online forums, blogs, and other social media where discourse occurs. The focus was on statements related to the purpose, content, and process elements of nonprofit performance measurement, and the goal was to include all expressions of perspectives on the topic of study (Webler et al., 2009). Statements were included when a clear opinion was disclosed. Ideal statements have excess meaning and can be interpreted differently by participants during the Q sort (Webler et al., 2009).

After generating a concourse thought to be inclusive and exhaustive of the opinions held about NPO performance measurement, a finalized subset of statements was selected for use in the Q sample (van Exel & de Graaf, 2005). The initial concourse

included more than three hundred statements. The Q sample is a representative list of statements derived from the concourse that has been condensed to reduce ambiguity and duplication yet still represents the main ideas, opinions, and views collected in the concourse, typically consisting of 40 to 50 statements (Akhtar-Danesh et al., 2008). “The nature of the stimuli making up the Q sample is constrained only by the domain of subjectivity in which the researcher is interested” (McKeown & Thomas, 2013, p. 12).

Q samples can be structured or unstructured in how statements are excluded from the concourse (McKeown & Thomas, 2013). “Structured sampling uses statements that try to offset potential weaknesses found in other statements, provide focus, and place boundaries on the topic of study” (Buchman, 2009, p. 55). A structured Q sample was created from strategic sampling of theoretical categories identified during the literature review (Webler et al., 2009). Researcher bias in the selection of statements comprising the Q sample is minimized through strategic and inductive sampling practices (Webler et al., 2009).

The strategic sampling approach draws upon theoretical categories observed in the concourse, derived through inductive analysis. “Inductive analysis works from the raw data to generate conceptual categories that have explanatory potential” (Webler et al., 2009, p. 14). For this study, statements were organized into purpose, process, and content categories from which 14 representative statements best representing each category were selected to formulate the Q sample. All selected Q sample statements were representative of nonprofit performance measurement, achieving homogeneity (S. R. Brown, 1980). Each Q sample statement represented a different thought within the

theoretical category it represents, achieving heterogeneity with respect to the concourse (S. R. Brown, 1980).

Q samples should be diverse so that participants will react both positively and negatively to the statements, thereby sorting half of the statements on each side of the forced quasi-normal distribution and creating a distensive zero (Webler et al., 2009). The theoretical categories with an equal number of statements in a structured sample prevent the sample from being weighted toward a single outcome and eliminate the need for the sample to be exhaustive (Stephenson, 1953; Webler et al., 2009). Although the theoretical categories help to achieve balance, variety, and diversity of ideas represented in the Q sample statements, these theoretical categories are not relevant to the results.

For this study, a group of three nonprofit leaders with experience in nonprofit performance measurement assisted in a review of the potential Q sample to assist with content and face validity of the statements (Akhtar-Danesh et al., 2008). Selected because of their professional or academic experience and knowledge of nonprofit performance measurement, the domain experts reviewed the Q sample statements to ensure homogeneity and heterogeneity were met. Their feedback provided direction for statement modifications, including removal of duplicate ideas, improving clarity and consistency, and ensuring maximum representativeness of ideas within the concourse. The finalized Q sample was created from the initial Q sample statements and was used by participants in the Q sort.

According to Akhtar-Danesh et al. (2008), “The objective in Q methodology is to be able to describe typical representations of different viewpoints rather than to find the proportion of individuals with specific viewpoints” (p. 763). There will ideally be more

observations (statements) than variables (Q sorts) in a given study (Webler et al., 2009). Typical Q studies result in two to five social perspectives and require four to six participants to define each perspective (Webler et al., 2009). A 3:1 ratio of statements to participants is most common, with the lowest ratio of statements to number of participants typically being 2:1 (Webler et al., 2009). However, there must be at least four or five individuals to define each factor, which may require more participants than this ratio (S. R. Brown, 1980). Because this study's Q sample includes 42 statements, the goal was to include 20-40 participants in the P-set. Only a relatively small number of participants was necessary because they are the means for measuring variables rather than representing a population (Sickler et al., 2006). Additional participants have little impact on the factor scores because Q methodology results are generalizable to the defined factor type rather than the population (S. R. Brown, 1980).

P-set participants should not be selected at random, but rather have a theoretical relevance to the study and a distinctive stance on the topic that will help define a factor (S. R. Brown, 1980). Knowledgeable individuals with well-formed opinions are the best participants, and a purposive chain referral sampling technique allows the researcher to recruit participants who meet the range of opinions identified in the concourse (Webler et al., 2009). "A factorially designed P-sample . . . makes an overt attempt to sample people of theoretical interest. It provides a degree of comprehensiveness not found in samples chosen solely on the basis of availability" (McKeown & Thomas, 2013, pp. 37-38).

This study's P-set were recruited through chain referral sampling to solicit a group of participants who are leaders in nonprofit management and nonprofit performance measurement. Much like snowball sampling, chain referral sampling

involves relying upon individuals with experience with the phenomenon of interest to provide a series of participant referrals (Penrod et al., 2003). The difference between snowball and chain referral sampling is that chain referral sampling allows for strategic access of multiple networks, which aids in the expansion of the investigation beyond a single social network (Penrod et al., 2003).

For the purpose of this study, nonprofit leaders were defined as (a) academic experts who have published within the field in the past 5 years or (b) professionals with 5 or more years of leadership experience within the field. Professionals are nonprofit senior management, evaluators, grantors and funders, consultants, and watchdog agency leaders. Members of nonprofit management and performance measurement groups or associations that focus on improving the nonprofit sector were identified as well.

Initial contact with potential participants occurred through telephone, e-mail, and U.S. mail. Potential participants were identified during the literature review and through research of charity watchdog agencies, foundations, and professional evaluators and nonprofit consultants. Contact also occurred through the researcher's personal and professional network. Those contacted were asked to participate if their experience meets the criteria for this study, determined by a brief demographic questionnaire. Participants were asked to introduce the researcher to other individuals who meet requisite experience requirements and who might have an interest in participating in the study, with an expressed interest for those who might have opinions different from their own.

After the Q sample and P-set were created and recruited, respondents were instructed to complete a Q sort, a process that should take no more than 1 hour. The Q

sort is the method by which information is collected for subsequent factor analysis. “Respondents, or the P-set, are asked to rank order the statements from their individual points of view according to some preference, judgment, or feeling about them, mostly using a quasi-normal distribution” (de Graaf & van Exel, 2008, p. 65). The Q sort is a reflection of the participants’ thinking, evaluation, and interpretation of the Q sample stimuli (S. R. Brown, 1980).

Participants were given a condition of instruction that iterates the context for interpretation and reaction to the Q-statements (Sickler et al., 2006) and were asked to read and rank order the Q sample statements into a Q sort table or score sheet (Akhtar-Danesh et al., 2008). For this study, the condition of instruction asked participants to sort the statements with reference to which characteristics, behaviors, and practices were most important and most unimportant to effective nonprofit organizational performance measurement. Before they assigned a rank order to the statements, participants divided statements into three piles: those they believe are most important to effective nonprofit organizational performance measurement; those they believe are most unimportant; and those they are neutral, doubtful, or undecided about (van Exel & de Graaf, 2005). This process allows for capturing the participants’ qualitative opinions and perceptions about the topic for further quantitative review and analysis.

### **Study Population and Sample**

A population is a group of individuals or organizations sharing a common characteristic, and a target population shares a common characteristic of interest to the researcher (Creswell, 2005). In the United States, there are 1.4 million tax-exempt 501(c)3 nonprofit organizations (Independent Sector, 2012a), each of which is led by an



executive director, chief executive, founder, or other leader, in addition to an unknown number of experts, researchers, evaluators, and consultants participating in the nonprofit organizational management and performance measurement dialogue. The sample selected from the target population will represent the entire population (Creswell, 2005).

As it would be cost and time prohibitive to study the entire population, the study population was limited to a portion of the population (Cooper & Schindler, 2008). Because “the objective in Q methodology is to be able to describe typical representations of different viewpoints rather than to find the proportion of individuals with specific viewpoints” (Akhtar-Danesh et al., 2008, p. 764), identifying the size of the population is unnecessary. A sample of nonprofit leaders who have experience in nonprofit management and performance measurement were the target population of this study. For the purposes of the study, nonprofit leaders were defined as (a) academic experts who have published within the field in the past 5 years or (b) professionals with 5 or more years of leadership experience within the field. Professionals included NPO senior management, evaluators, grantors and funders, consultants, and watchdog agency leaders.

There are many methods for determining a sample of the population to research. Sampling techniques are categorized into probability and nonprobability methods, and for an exploratory study with limited time and budget, a nonprobability method is appropriate (Cooper & Schindler, 2008). The nonprobability sampling technique selected for the study is purposive judgment. Purposive judgment sampling “occurs when a researcher selects sample members to conform to some criterion” (Cooper & Schindler, 2008, p. 197) and is an appropriate technique for the exploratory and subjective nature of this study. In purposeful sampling, the researcher establishes criteria for inclusion or

exclusion based upon the research questions and the participant's ability to answer questions about a predetermined topic (Breckenridge & Jones, 2009). Chain referral sampling, or using participants and informants to identify potential participants for the study (Gelo, Braakmann, & Benetka, 2008), is a sampling frame that can further enable the researcher to gain access to nonprofit leaders with experience in nonprofit management and performance measurement.

While nonprobability samples, and more specifically snowball and chain referral samples, raise questions of representativeness, the procedure is appropriate for exploratory purposes (Rubin & Babbie, 2010), especially in a population that is hidden or for which total membership is unknown. "Although probabilistic representative samples are impossible in studies of hidden populations, non-probability designs can be set up for making at least *well-founded* inferences about the population of interest" (Spren & Zwaagstra, 1994, p. 478). A characteristic and strength of Q is that the number of people who hold common views is not a concern but rather the number of common perspectives that exist, sampling a "population of viewpoints rather than a population of people" (Jedeloo, van Staa, Latour, & van Exel, 2010, p. 595).

Because Q researchers focus on the spectrum of viewpoints rather than representativeness of the population, small sample size and bias are not relevant (Akhtar-Danesh et al., 2009). A growing body of Q studies has affirmed that two to five social perspectives about a given topic are typically identified, and only four or five participants are required to define a factor (perspective; Webler et al., 2009). Participants in this study were specifically asked to refer those in the field with a known difference in

opinion or perspective in an attempt to reduce bias and increase the ability to find different social perspectives held by nonprofit leaders.

A danger in snowball sampling is the possibility of tapping into a single or limited number of networks, which can limit the application of findings in studies concerned with representativeness (Penrod et al., 2003). However, in this study, representativeness was not as pivotal as locating and identifying as many viewpoints as possible. Chain referral sampling is similar to snowball sampling, relying on participant referrals to study the phenomenon of interest, “however, *multiple networks are strategically assessed to expand the scope of investigation beyond one social network*” (Penrod et al., 2003, p. 102). This study included nonprofit leaders across the United States, especially targeting those holding different roles within the sector, thereby increasing the potential to identify a wide range of differences in opinion.

As previously mentioned, potential participants were identified during the literature review and through research of charity watchdog agencies, foundations, and professional evaluators and nonprofit consultants. Contact also took place through the researcher’s personal and professional network. Participants were asked to introduce the researcher to other nonprofit leaders who fit the sample criterion for this study and who might hold different viewpoints of the topic under study. The goal of the study was to include 20-40 participants, and the final participant group included 22 nonprofit leaders.

### **Informed Consent**

Each NPO leader identified as a potential participant received a letter of invitation to participate in the study (see Appendix A), followed by a telephone call or e-mail. The letters included an informed consent statement for the study (see Appendix B), which

outlined the purpose of the study, its duration, and procedures for completion. The letter informed participants of confidentiality and the level of risk associated with the study, explained voluntary participation, and noted the ability to withdraw from participation. The study has no known associated risks any greater than risks experienced in daily life. No direct benefit for participation in the study was expected. Each participant was offered an executive summary of the results following the completion of the study.

Individuals who agreed to participate in the study were asked to mail the original signed consent form to the researcher in the self-addressed stamped envelope provided via U.S. mail. Consent forms and all written or printed study documentation was collected and stored separately from research information and kept in a secure, locked file cabinet in the home of the researcher. The researcher is the only person with a key and access to the consent forms and the research data. All electronic research data were stored on a password-protected external hard drive that is also stored in a separate locked cabinet in the home of the researcher. All research materials will be kept for 5 years and then destroyed by secure deletion and document shredding.

### **Confidentiality**

Upon completing the informed consent form, participants received a unique random identifier rather than using their names or organizational affiliations to ensure confidentiality. “Confidentiality is maintained when anything that is learned about the participant is held in the strictest of confidence” (Salkind, 2008, p. 64). In addition to the assignment of unique identifiers to each participant, a single master list of participants and identifiers was maintained (Salkind, 2008). Only the researcher has access to this list. The unique identifiers were used in all subsequent data collection and analysis

documentation, thereby ensuring participant confidentiality. The list containing identifiers and all raw data files will remain in separate and secure locations within the researcher's home to reduce the likelihood of linking specific data to study participants, to maintain participant anonymity, and to ensure confidentiality. Although the results of this study may be published, at no time will the identity of study participants be linked to published information, and only aggregated results will be reported.

### **Geographic Location**

Participants in the study were limited only by their expertise in nonprofit management and performance measurement, as previously defined, and by their residence in the United States. The goal of the study was to identify the subjective perceptions of nonprofit performance measurement held by nonprofit leaders in the United States, and the purposive chain referral sampling method encouraged participants to refer other nonprofit leaders who meet the sampling criteria. The discourse regarding nonprofit performance measurement is occurring on a national level. Sampling at a national level enabled the identification of American social perspectives regarding effective nonprofit performance measurement.

### **Instrumentation**

Participants in a Q methodology study project their subjective meaning upon statements during the Q sort, and each participant's opinions and perceptions create meanings different from other participants. As such, validity in Q methodology is determined differently from other methodologies. "Validity of a Q-study is evaluated by content, face, and Q sorting validity" (Akhtar-Danesh et al., 2008, p. 766). Content is validated through the literature review and by a team of domain experts, a common

qualitative research practice, in addition to pilot studies or field tests (Akhtar-Danesh et al., 2008). An external criterion for evaluating an individual's response to statements does not exist, so the set of rank-ordered statements serves as a valid expression of the individual's opinion (Akhtar-Danesh et al., 2008; S. R. Brown, 1980).

A small panel of nonprofit leaders allowed for further refinement of the Q sample. Three domain experts were recruited for the panel and completed a review of concourse statements. Domain experts were identified by the same criterion chosen for the sample, and the same identification, solicitation, and recruitment process was used to identify the panel. The group of nonprofit leaders helped to determine the opinions that were omitted, underrepresented, or incorrectly represented within the Q sample, in addition to ensuring that homogeneity and heterogeneity were met in the statements included. The panel guided further refinement of Q sample statements and ultimately resulted in the final Q sample for this study.

### **Data Collection**

Participants in this study were nonprofit leaders with the requisite experience who return a signed consent form. Participants received Q sort materials through the U.S. mail. Each packet of materials consisted of a standardized list of instructions for completing the Q sort, a list of 42 alphabetized statements, Q sort index cards, and a short demographic questionnaire (see Appendices C, D, and E). Q sort index cards contained the 42 statements numbered from 1 to 42.

Participants were instructed to rank Q sample statements from most unimportant to most important to NPO performance measurement (e.g. -5 to +5). To begin this process, participants were instructed to sort the statements into three piles. One pile was

statements the participant found to be most important to nonprofit performance measurement. Another pile was those statements the participant found to be most unimportant to nonprofit performance measurement. The third pile included the remaining statements.

The standardized instructions then told participants to rank order the statements into the range of most important characteristic (+5) to most unimportant characteristic (-5) for nonprofit performance measurement using the Q sort matrix as a guide (see Figure 1). The list of alphabetical statements included a line next to each statement for the participant to record the scores upon completion of the Q sort. A Q sort distribution that provides a forced-choice condition requiring participants to match Q sort cards on a preset Q sort matrix with one space per card was included (see Figure 1).

	Most Unimportant				Neutral				Most Important			
Ranks	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	
Items	2	3	4	4	5	6	5	4	4	3	2	
	41	38	34	30	25	19	14	10	6	3	1	
	42	39	35	31	26	20	15	11	7	4	2	
		40	36	32	27	21	16	12	8	5		
			37	33	28	22	17	13	9			
					29	23	18					
						24						

Figure 1. Quasi-normal Q sort matrix.

Upon completion of the Q sort and subsequent recording procedures, participants returned Q sort materials in the provided self-addressed stamped envelope through the U.S. mail to the researcher. Participants who did not respond within 2 weeks of receiving the materials received a follow-up telephone call from the researcher.

### Data Analysis

The data were analyzed to determine whether statistically significant factors from the individual perspectives exist, factors discernible only to nonprofit leaders experiencing the phenomenon under investigation (Stephenson, 1953). Q methodology includes three main sets of statistical procedures: correlation, factor analysis, and factor score computation (McKeown & Thomas, 2013). Participants bestowed subjective meaning to the Q-set, a sample of statements about performance evaluation, which were correlated between participant profiles for similar viewpoints (van Exel & de Graaf,



2005). Analysis uncovered patterns from individual responses, which were composed into social perspectives (Sickler et al., 2006).

Q sorts represent modeled viewpoints from which data analysis is conducted. Intercorrelation of the variables, or the  $N$  Q sorts, and factor analysis of the  $N \times N$  correlation matrix are the primary means of data analysis (McKeown & Thomas, 2013). Persons, rather than traits or statements, are correlated, resulting in factors representing points-of-view or social perspectives, and the participants' association with the identified perspectives is indicated by the participants' loading on each factor (Akhtar-Danesh et al., 2008; McKeown & Thomas, 2013). The Q sort patterns identified through inverted factor analysis are the basis for composing social perspectives (Sickler et al., 2006, p. 356).

PQMethod 2.33 software, developed by Schmolck and freely available free on the PQMethod website, is the analysis software most commonly used by Q researchers (Valaitis et al., 2007). PQMethod 2.33 software serves to extract factors and determine data patterns (Schmolck, 2002). Factors can be theoretically significant even if they are not statistically significant (McKeown & Thomas, 2013).

After the data are analyzed, the findings, or factor scores of each statement within each factor are interpreted based upon the statements that were most relevant to the factor, known as the distinguishing statements (Akhtar-Danesh et al., 2008). "The validity of the interpretation relies on the use of factor analysis in extracting the distinguishing statements and the use of domain experts in interpretation" (Akhtar-Danesh et al., 2008, p. 68). The interpretation is further validated through follow-up comments about nonprofit performance measurement collected from participants who

comprise each factor during the post sort survey and follow-up phone calls if necessary (Akhtar-Danesh et al., 2008).

The datum was interpreted based upon concepts recurring throughout the factor arrays. Statements were interpreted holistically, their unique concepts applied within the greater context of the factor array (Watts & Stenner, 2012). The viewpoints were organized and categorized by the concepts found within the distinguishing statements. Further, the concepts were examined for relationships between the statements and the theoretical framework.

### **Summary**

This Q methodology study involved an attempt to balance the strengths of qualitative inquiry of subjectivity with quantitative factor analysis to determine the various subjective social perspectives held by American nonprofit leaders concerning nonprofit performance measurement. Q methodology was the appropriate methodology as it “encompasses a distinctive set of psychometric and operational principles that, when conjoined with specialized statistical applications of correlational and factor-analytical techniques, provides researchers a systematic and rigorously quantitative means for examining human subjectivity” (McKeown & Thomas, 2013, p. 7). Q methodology transforms subjective statements into objective outcomes in a feasible and effective manner (Akhtar-Danesh et al., 2008). Because the goal of Q methodology is not to determine representativeness of a population of people, this study was completed by small number of well-selected participants (Stephenson, 1953). The study aimed to recruit 20-40 nonprofit leaders in the United States with experience in nonprofit management and performance measurement acquired through a chain referral sampling

method. The goal of the study was to identify the common factors, or social perspectives, held by nonprofit leaders about nonprofit performance measurement and how these perspectives might affect the development of a performance measurement model to replace financial efficiency. Chapter 4 will include a discussion on the data analysis and research findings.

## Chapter 4

### Results

The purpose of the Q methodology study was to gain a better understanding of the subjective perspectives of nonprofit leaders in the United States concerning effective organizational performance measurement. Understanding and identifying the elements that each social perspective considers most important in nonprofit performance measurement could facilitate the creation of a model that has broad application across the diverse nonprofit sector. This chapter includes the findings according to the research questions described in Chapter 3. The chapter contains a description of the factor structure, followed by the interpretation of the factors according to the perspectives of 22 nonprofit leaders in the United States.

Q methodology includes three main sets of statistical procedures, correlation, factor analysis, and factor score computation, and qualitative interpretation based on these data (McKeown & Thomas, 2013). Participants bestowed subjective meaning to the Q-set, a sample of statements about performance measurement correlated between participant profiles for similar viewpoints (van Exel & de Graaf, 2005). Analysis uncovered patterns from individual responses that were composed into social perspectives (Sickler et al., 2006). Q sorts represent modeled viewpoints from which data analysis took place. Chapter 4 includes detailed analysis of the Q sorts, descriptive demographic data, and clarification gained from postsort interviews in relation to the research questions and the resulting factors.

## **Overview of Data Collection**

Recruitment of participants began on November 30, 2012, and concluded on March 9, 2013. The recruitment process involved making contact with 135 nonprofit leaders. A sample of nonprofit leaders who have experience in nonprofit management and performance measurement were the target population of this study. For the purposes of the study, nonprofit leaders were (a) academic experts who have published within the field in the past 5 years or (b) professionals with 5 or more years of leadership experience within the field. Professionals include nonprofit senior management, evaluators, grantors and funders, consultants, and watchdog agency leaders.

Individuals identified through the literature review included academics, authors, nonprofit leaders, charity watchdog leaders, and consultants who had been vocal in the ongoing nonprofit performance measurement dialogue. Identified through chain referral sampling, participants in the study received a specific request to refer those in the field with a known difference in opinion from their own in an attempt to find different social perspectives held by nonprofit leaders (Penrod et al., 2003). Contact occurred primarily through e-mail, although some contact was by phone.

Thirty-four individuals received recruitment packets through the U.S. mail. Each packet of materials consisted of a standardized list of instructions for completing the Q sort, a list of 42 alphabetized and randomly numbered statements printed on index cards, and a short demographic questionnaire (see Appendices C, D, and E). Follow-up by phone and e-mail occurred 2 weeks after mailing the packets to ensure participants received the materials and to answer any questions. Of the 34 who initially expressed interest in the study, 22 returned completed informed consent forms, Q sort matrices, and

postsort questionnaires. Of the 12 who originally agreed to participate but did not return materials, six withdrew because of time constraints, three withdrew because they misunderstood the nature of the study, and three were nonresponsive to follow-up attempts.

Participants received a condition of instruction that provided the context for reaction to the Q-statements (see Appendix D). The condition of instruction was as follows: “Sort the statements according to the characteristics, behaviors, and practices that promote effective nonprofit organizational performance measurement.” Instructions contained a request that participants would read and rank order the 42 Q sample statements (see Appendix C) into a Q sort matrix or score sheet ranked from those statements they believed were most unimportant to those they believed were most important to effective nonprofit performance measurement (e.g., -5 to +5). After participants completed the Q sort exercise, they received a request to fill out the postsort questionnaire (see Appendix E), which included questions about the exercise in addition to demographic information. All participants who returned completed packets received a handwritten thank-you card in the mail and an executive summary of the results.

### **Description of Participants**

Participants completed a demographic profile with a variety of information about themselves, in addition to questions about their nonprofit organization, including organization type, location, operational budget, number of full-time staff, and the role of performance measurement within their organization. The demographic information for the total participant group appears in Figures F1-F11 in Appendix F. Discussion of the demographics for each factor are in the factor review section, although no patterns in

demographic responses appeared by factor. Resulting viewpoints did not cluster participants by years of experience, sector role, education, or any of the other five demographic categories.

### **Data Analysis**

Correlation is the first step in statistical analysis. PQMethod 2.33 software was suitable for analyzing the data (Schmolck, 2002). Individual Q sort matrices correlate to all participant sorts, which indicates the level of agreement between sorts and similarities in perspectives (van Excel & de Graff, 2005). Individuals' Q sorts are the variables in Q method. The correlation matrix generated has correlations ranging from -1.0 to +1.0 based upon the relationships between variables (Webler et al., 2009). Principal component analysis of the correlation matrix revealed the factor structure. The analysis involved rotating the resulting components orthogonally using varimax rotation and then subjecting them to a final Q analysis that used a calculation of *z* scores for all the original Q sample statements occurring within each factor to differentiate the factors (Giannoulis, Botetzagias & Skanavis, 2010).

Factor loadings range from -1.0 to +1.0 and represent the variation or commonality within the factor (Giannoulis et al., 2010).

Persons significantly associated with a given factor, therefore, are assumed to share a common perspective. . . . Therefore, each respondent's factor "loading" indicates the degree of association between that person's individual Q sort and the underlying composite attitude or perspective on that factor. (McKeown & Thomas, 2013, p. 17)

Q sorts with a very high loading on only one factor are pure, whereas Q sorts loading on

more than one factor are confounded, representing a hybrid perspective (Giannoulis et al., 2010). This study did not involve using confounded sorts as defining sorts for the factors. Nonsignificant sorts do not achieve a significant load and are those that are not similar to any of the resulting perspectives (Webler et al., 2009). Q sorts with loadings that reach significance on only one component define factors and help to produce a single set of factor scores displayed in a factor array (McKeown, Hinks, Stowell-Smith, Mercer, & Forster, 1999).

The formula for calculating factor loading significance is  $\alpha < .01$ ; significance =  $2.58(1/\sqrt{N})$ , with  $N$  equaling the number of Q-statements (McKeown & Thomas, 2013). The significance level of  $\alpha < .01$  was suitable because it is more rigorous than  $\alpha < .05$  and further limits opportunities for consensus while ensuring statistical significance for defining sorts (Sickler et al., 2006). Significance for the study was .398 and rounded to .40, where  $N = 42$  and  $\alpha < .01$ . Further increasing significance to .45 resulted in a more rigorous and clearer definition of differences among factors while minimizing the number of confounded sorts.

A three-factor solution was the best solution after several test rotations. Sixteen of the 22 sorts loaded significantly on only one of the three factors. Factor 1 had eight significant sorts, while Factors 2 and 3 both had four significant sorts. Four sorts were confounded and two sorts were nonsignificant. Table 2 contains the factor matrix. To ensure confidentiality, participants received a coded signifier.



Table 2

*Factor Matrix With Defining Sorts*

QSORT	Factor 1	Factor 2	Factor 3
P18	<b>0.8373x</b>	-0.0306	0.1373
P21	<b>0.7187x</b>	0.1585	0.2726
P16	<b>0.7122x</b>	0.1121	0.2329
P04	<b>0.7078x</b>	0.2112	0.2261
P14	<b>0.6655x</b>	0.0577	-0.1571
P07	<b>0.5661x</b>	0.3597	0.3158
P09	<b>0.5453x</b>	0.3657	0.3158
P12	<b>0.4695x</b>	0.3301	0.4040
P10	-0.0834	<b>0.7513x</b>	0.1357
P03	0.2001	<b>0.7210x</b>	-0.0856
P15	0.2075	<b>0.6775x</b>	0.2824
P20	0.4129	<b>0.6296x</b>	0.1238
P01	0.1194	0.0152	<b>0.8062x</b>
P05	0.3061	0.1574	<b>0.6841x</b>
P02	0.3528	0.3370	<b>0.5089x</b>
P06	-0.0946	0.0209	<b>0.5043x</b>
P11	0.4654	0.5046	-0.0740 (confounded)
P13	0.0690	0.5750	0.4566 (confounded)
P19	0.5598	0.4954	0.1135 (confounded)
P22	0.4635	0.2589	0.4844 (confounded)
P08	0.2860	0.3892	0.2229 (nonsignificant)
P17	0.3682	0.1561	0.0332 (nonsignificant)
Number of sorts	8	4	4

*Note.* Numbers with bold font and an “x” represent defining sorts for each factor.

Table 3 contains the correlations of the factor scores of statements between factors. The correlations ranging from .3306 to .4155 explain the similarities between the pair’s factor arrays (Giannoulis et al., 2010). Although there are commonalities between the factors, there is enough difference to represent the individual views of effective nonprofit organizational performance measurement. The basis for examining the three resulting factors was the entire array of *z* scores or factor scores for the statements for that factor, distinguishing statements, especially those statements deemed most important (ranked scores of +3 through +5) and most unimportant (ranked scores of -3 through -5)

by each factor. These scores describe the beliefs of each factor, in addition to clarification gained through the postsort questionnaire and follow-up interviews.

Table 3

*Correlations Between Pairs of Factors*

Factors	1	2	3
1	1.0000		
2	.4145	1.0000	
3	.4155	.3306	1.0000

**Interpretation of Factors**

Using results from the Q methodology analysis, the following sections address the research questions as they pertain to the three-factor solution. “Each factor is interpreted using the statements specific to that factor. In the lexicon of Q methodology, these statements are called distinguishing statements because they define the uniqueness of each factor” (Akhtar-Danesh et al., 2008, p. 768). Researchers typically use two or three columns of the highest positive and negative ranked statements for interpretation of the factors, in addition to comments and demographic information provided by participants in the postsort questionnaire (Dziopa & Ahern, 2011). Follow-up interviews with participants helped to clarify the resulting viewpoints.

The data were interpreted based upon concepts recurring throughout the factor arrays. Statements were interpreted holistically, their unique concepts applied within the greater context of the factor array (Watts & Stenner, 2012). The viewpoints were organized and categorized by the concepts found within the distinguishing statements. Further, the concepts were examined for relationships between the statements and the theoretical framework.

The three viewpoints named were Road Map, Management Tool, and Weakest Link, and explorations follow in the succeeding sections, organized by research question.

**Research Question 1.** The first research question was as follows: What do nonprofit leaders believe to be the most important elements of effective nonprofit organizational performance measurement? The individuals on each factor had differing opinions about the characteristics, behaviors, and practices; however, there were some common themes throughout the viewpoints.

First, performance management was the focus of all perspectives. Consistent with all three viewpoints, performance measurement cannot stand alone to be effective. Performance measurement is a tool in the performance management system that should help nonprofit leaders to learn and improve performance. Performance management is the larger system for effective performance, with performance measurement being one of the system's practices.

The second theme was the unfunded mandate for performance evaluation. The participants called attention to how funders expect performance data as a condition of the funding, yet few grantors include funding for administrative and evaluation purposes. The participants also conveyed a desire for a sense of partnership with their funders, both working in concert toward the same goal of better service delivery for beneficiaries. Characteristics of an effective performance measurement model are partnership and shared goals.

The final theme was outcome measurement. The participants favored outcome measurement as the most effective practice for evaluating a nonprofit's performance. When nonprofit leaders have a better understanding of their outcomes and performance

data, then there is opportunity for greater collaboration, learning, and increasing the sector’s collective impact. An explanation of each factor appears in greater detail below.

**Factor 1: Road Map.** Eight sorts defined Factor 1, Road Map. The individuals on this factor described performance measurement as a means to improve the organization and stimulate learning. Results provided important information, serving as a formative tool and road map to guide improvement. The members of Road Map desired to serve their beneficiaries better and achieve mission-based goals. Provided in Table 4 are the nine most important and the nine most unimportant statements for the Road Map factor, along with each statement’s array position and *z* score. Table G1 with complete statements is located in Appendix G.

Table 4

*Most Important and Most Unimportant Statements and Scores for Road Map Condensed*

Item Number	Statement	Array Position	<i>z</i> -Score
33	Performance measurement should be at the organizational...	5	1.66*
3	A new performance measurement model should help...	5	1.60
28	NPO performance measurement metrics must never...	4	1.59*
7	An effective performance measurement model should...	4	1.53
35	Rather than asking traditional strategic planning...	4	1.43
9	An effective performance measurement model will...	3	1.10
4	A performance measurement system that simplifies the...	3	0.98*
34	Performance measurement should begin with the end in...	3	0.89
24	NPO capacity, or lack thereof, impacts effective...	3	0.88
15	Donors will have a more complete picture of NPOs...	-3	-0.95
40	The sole purpose of NPO performance measurement is...	-3	-1.03
41	Too many NPOs are concerned with sustaining...	-3	-1.08*
36	Roughly 3,000 charities garner 60% of all revenue that...	-3	-1.13
14	Donors need comparative data and ratings providing...	-4	-1.49*
23	Most NPO leaders don’t communicate results or lack...	-4	-1.60*
32	Performance measurement methods are forced upon...	-4	-1.63
31	Performance data should stand on its own, minimizing...	-5	-2.01
17	Financial efficiency will always be the most important...	-5	-2.14

*Note:* Numbers with asterisks denote distinguishing statements

Of the eight participants who established Road Map, six were women and two were men. Members of the Road Map group were highly educated, as five had earned master's degrees and three held doctorates. Five had more than 20 years of experience in the nonprofit sector, two others had 15-19 years, and one had between 10 and 14 years of experience. Half of the participants were executive directors or CEOs of a nonprofit organization, and two others held senior management positions. Individuals on this factor included a foundation professional and an academic, and the factor accounted for 23% of the variance.

According to members of Road Map, the primary purpose of performance measurement is to promote performance improvement (Statement 3: +5). An effective organization is one that addresses and adapts to the ever-changing needs of its community (Statement 35: +4). Community needs should be the basis for the organization's mission and goals. For this viewpoint mission attainment is paramount, and above all else, performance measurement should guide the nonprofit to fulfilling its mission (Statements 33: +5; 28: +4; 41: -3). Effective performance measurement is an organization specific and mission-focused process (Statements 33: +5; 7: +4).

For the members of Road Map, performance measurement is a form of formative assessment that allows organizational leaders to understand where an organization rates in its path to mission attainment. Performance measurement should not be viewed as a static or summative procedure conducted for the benefit of external parties (Statement 33: +5). Rather, performance measurement should serve as a progress check on the road to goal attainment (Statements 9: +3; 5: +2). In this viewpoint, performance evaluation

refers to developmental evaluation, which facilitates the search for both the ends and the means.

The focus of performance measurement should be on helping nonprofit leaders attain the organization's mission and fulfill community needs (Statement 35: +4). Performance measurement should not hinder the organization's progress on the mission. Rather, it would be complementary to the organization's operations (Statement 28: +4). Performance data would be used to correct the course of the organization. Performance measurement is woven into the fabric of the organization, serving as a formative and developmental practice to inform performance improvement (Statement 33: +5). As a dynamic tool it allows the organization to experiment and innovate (Statement 33: +5). The organization has a clear vision and performance measurement serves as a guide to reach that destination.

Additionally, the purpose of performance measurement is not to appease donor demand for information or to demonstrate impact or results. Performance data is not for reporting purposes or communicating impact (Statement 18: +1). Standardization of performance metrics and externally conducted evaluation are dismissed in favor of performance measurement completed by members of the organization based upon stated goals (Statements 37: +1; 14: -4; 23: -4). This is in part due to a lack of confidence that standardized performance data will lead to an increase in donations or actionable data to improve performance (Statements 15: -3; 37: +1). These leaders believe that the nonprofit sector is responsible for important work that is not easily quantified. Standardized evaluation would likely measure that which is easy to quantify rather than true performance improvement indicators. Members of Road Map identify the gap in the

discourse stems around the design of an effective performance measurement method (Statements (29: +2; 27: +1).

Road Map participants believed stakeholders deserve performance information and valued their input, which showed awareness of their resource dependencies. However, these leaders preferred not to institutionalize expectations that prohibit or affect mission progress. Further clarified in the postsort questionnaire, the proposition of a standardized performance metric to evaluate and refine strategy is appreciated, however, the plausibility is questioned from a total cost standpoint. The wide difference in communities and constituencies prevents standardization as a realistic or attainable goal for the sector (Statement 37: +1). Such standardization would likely prevent nonprofit leaders from recognizing the subtle nuances that are specific to addressing their communities' needs. Despite collecting performance data for performance improvement, the reality of resource dependency creates awareness that this data can and will be used by stakeholders for accountability purposes. Further, both positive and negative findings will reflect upon organizational leaders. Performance measurement should not create fear of retribution but rather should produce data that will provide insights and spur changes addressing those problem areas.

Road Map participants welcome expert analysis and recommendations for how to improve performance. However, the members of this viewpoint frowned upon performance measurement solely conducted by external parties. Road Map participants recognized the dangers of not including outside performance review, namely bias and the tendency to maintain the status quo (Statement 31: -5). Organizational leaders should conduct performance measurement with an unbiased outside party to review the

evaluation process. While the standard should be to conduct internal performance measurement, nonprofit leaders should seek external feedback (Statement 31: -5). External parties have expertise, resources, and insights that would strengthen the organization's performance measurement process thereby advancing the organization's ability to continuously improve (Statements 24: +3; 22: +2).

Road Map participants expressed concern regarding the cost and motivation to conduct performance measurement (Statements 13: +1; 15: -3). Although members of Road Map believed performance measurement is important, they highlighted the externally driven mandate for performance data from funding bodies that is neither funded nor conducive to mission progress (Statement 37: +1). The expectations of funders, each with their own guidelines and data recommendations, cannot isolate performance insights (Statement 27: +1).

Participants sharing the Road Map viewpoint emphasized effective nonprofit performance measurement takes place at the organizational level. Measures should be easy to implement, cost efficient, and dynamic metrics for assessing organizational performance, inspiring process innovation, and allowing for experimentation (Statement 33: +5). Measures should be motivating for the organization's members, as well as easy to use. All nonprofit leaders should conduct performance measurement and create metrics with the desired end-state in mind. Organizational leaders must participate for organizational learning and performance management to occur. Additionally, in those organizations that need it, building performance measurement capacity should involve training, adding infrastructure, and so forth through partnerships with and investments from donors (Statement 24: +3).



According to the members of the Road Map viewpoint, financial efficiency is not the most important measure and should not be the sole measure of organizational performance (Statement 17: -5). Financial health is important, yet financial health does not guarantee positive outcomes. Road Map participants believed measures should be specific to the organization measured, should not be oversimplified for the sake of standardization, and should include multiple dimensions of performance (Statement 7: +4; 4: +3). Supporters of the viewpoint believed community need and mission attainment need to come first, but poor financial health will prevent the attainment of goals. Members of this perspective value their part in the open system, yet seek to achieve greater balance in power dynamics.

Members of the Road Map viewpoint did not believe all organizations have the same issues preventing effective performance. This viewpoint disagreed with statements that define specific metrics, measures, or high performance, in deference to leaving that decision to the leaders of the specific organization. Outside of the focus on performance measurement as a guide to address community needs and achieve the organizational mission, statements that are prescriptive of decisions that should be made internally are dismissed. The purpose of a new performance measurement model is not to identify unbalanced power dynamics, undercapitalization, or to blame nonprofit leaders for lack of discipline or self-accountability (Statements 1: -1; 25:0). Questions about nonprofit organizational productivity, culture, and measurement characteristics are not prioritized (Statements 21: 0; 12: 0; 39: 0). Each of these elements is organizational specific and broad generalizations about nonprofit performance are avoided.

These nonprofit leaders are careful not to prescribe a universal behavior or decision for other organizations, deferring to the leadership of each specific organization to decide what is best for their path to mission attainment. Likely, this is due to their own frustration when assumptions are made about the organizations they lead, their own wisdom and community awareness negated by sweeping generalities and exclusion from the important decisions that affect their organizations. When performance measurement is personal and specific to the organization and its members, it is more likely to produce a positive impact on performance.

***Factor 2: Management Tool.*** Four sorts defined Factor 2, named Management Tool. Participants of this factor defined performance measurement as a decision-making tool that generates evidence based dialogue with stakeholders. Resource dependencies and institutionalized expectations greatly influence the members of Management Tool's opinion and understanding of performance measurement. However, a genuine partnership that leads to performance improvement is the primary goal for Management Tool. Provided in Table 5 are the nine most important and the nine most unimportant statements for the Management Tool factor, along with each statement's array position and z score. Table G2 with complete statements is located in Appendix G.

Table 5

*Most Important and Most Unimportant Statements and Scores for Management Tool Condensed*

Item Number	Statement	Array Position	z-Score
8	An effective performance measurement model will...	5	1.44*
34	Performance measurement should begin with the end in...	5	1.42
3	A new performance measurement model should help...	4	1.39
21	Measuring and using outcomes data is essential for...	4	1.29
18	If NPOs want to prove legitimacy, they need to uphold...	4	1.28
41	Too many NPOs are concerned with sustaining...	3	1.23*
7	An effective performance measurement model should...	3	0.99
25	NPO leaders must commit to self-discipline and...	3	0.97
13	Devoid in the performance measurement discussion is...	3	0.94
6	A substantial flaw of many performance metrics is that...	-3	-0.90
30	People are obsessed with measurements; we want to...	-3	-0.94*
17	Financial efficiency will always be the most important...	-3	-1.07*
4	A performance measurement system that simplifies the...	-3	-1.08
40	The sole purpose of NPO performance measurement is...	-4	-1.09
42	We must reject overhead expense ratios as relevant...	-4	-1.19
32	Performance measurement methods are forced upon...	-4	-1.51
29	Outcome assessment is a highly complicated, uncertain...	-5	-1.94*
31	Performance data should stand on its own, minimizing ...	-5	-2.40

*Note:* Numbers with asterisks denote distinguishing statements

Of the four participants comprising Management Tool, three were men and one was a woman. This factor included three participants with doctorates, and in terms of experience within the field, two had 10-19 years' and one had more than 25 years' experience (one participant did not respond to these questions). Two participants were the executive director or CEO of their nonprofit organization, one was a foundation professional, and the last was a consultant. Members of Management Tool accounted for 16% of the variance.

For Management Tool participants, the perception of the external expectation for performance measurement is, in most cases, perceived as a genuine funder interest in performance improvement. In the context of this relationship, performance data would

help prove legitimacy to the funding organization, provided that both agencies share the desire to advance mission progress (Statement 18: +4). Members of Management Tool were cognizant that resource dependencies include institutional expectations. For these nonprofit leaders performance reporting is merely a facet of that relationship (Statements 14: +2; 15: +1).

According to Management Tool participants, the purpose of performance measurement is not for holding organizational leaders accountable but rather for promoting performance improvement (Statement 3: +4). Nonprofit leaders should use performance data to communicate to funders; however, communicating social impact is not the sole purpose of performance measurement (Statement 40: -4). While proving legitimacy to multiple constituencies is important, these NPO leaders desire evidence-based dialogue about organizational performance and goal attainment (Statements 18: +4; 34: +5). The individuals on this factor expressed a desire for partnership with funders, including mentorship, assistance interpreting performance data, and support in making performance improvements (Statement 31: -5).

Members of Management Tool did not believe the purpose of performance measurement is to reject overhead ratios or to demonstrate impact or results but rather an acceptable donor expectation (Statements 42: -4; 40: -4; 14: +2). This viewpoint discredits the need to question externally driven demand for performance data (Statements 1: -1; 30, -3; 23:0). The purpose of performance measurement is not to buck the institutionalized expectations of their resource dependencies but rather to work in concert with these expectations to develop a positive working relationship with funders.

Management Tool participants strongly reacted to the idea that performance data does not warrant expert analysis or recommendations (Statement 31: -5). According to this viewpoint, performance information should be subject to expert analysis and recommendations, further reiterating the symbiotic nature of relationships with donors. Participants sharing this viewpoint did not support considering performance measurement to be an overhead expense. Instead, performance measurement should be part of a grant or donation (Statement 13: +3). If donors want performance measurement, then they should be willing to fund this endeavor.

Members of Management Tool believed nonprofit leaders should create performance measurement with the end-state in mind and identify the mission-driven intended impact organizational leaders seek to achieve (Statement 34: +5). Performance measurement should not occur in isolation or solely as an internal evaluation process (Statement 16: -1). Organizational leaders should use insights from performance measurement to identify areas for improvement and performance management as an ongoing system to continuously improve performance (Statement 34: +5; 3: +4). Performance management, according to Management Tool participants' postsort questionnaires, is the most critical organizational practice. In organizations with active performance management, performance measurement is one of many tools used by management to improve organizational outcomes continually (Statement 25: +3).

Outcome measurement is not as difficult as it might seem; according to members of Management Tool, it is possible for all nonprofit organizations to implement and is easier to understand than most nonprofit leaders believe (Statement 29: -5). Outcome measurement is a method to generate organizational learning and performance

improvement (Statement 21: +4). Management Tool participants desired more discussion about why few nonprofit leaders use outcome measurement and speculated the language used can make performance measurement seem complicated and confusing for organizational leaders. Collaboration with funding partners could help minimize this confusion.

This perspective considers intended impact the most important performance metric, which is an organization-specific, mission-driven measure (Statement 8: +5). Members of the Management Tool perspective believed the organization's ability to make and communicate mission-driven progress and organizational outcomes is of utmost importance. The intent of this measure would be to improve the organization's ability to achieve desired outcomes (Statement 21: +4). Funders and grantees should have aligned goals, rather than competing goals, for social outcomes (Statement 11: +1). Nonprofit leaders should not chase funds but rather seek donors who support their intended impact (Statement 8: +5).

Management Tool members strongly believed measures should be specific to the organization measured and should focus solely on mission-based outcome achievement rather than multiple dimensions of performance (Statements 4:-3; 8: +5). These participants did not view financial efficiency as more important than the organization's ability to measure and communicate outcomes (Statement 17: -3). Further, management capacity, systems, and structures should not be the focus of performance measurement, and community benefit should not take precedence over definitions of high-performing organizations (Statements 41: +3; 35: 0; 6: -3). The goal should be to determine how to

solve challenging societal problems rather than generating performance data solely for reporting's sake.

**Factor 3: Weakest link.** Four sorts defined Factor 3, named Weakest Link. The individuals loading on this factor expressed that performance measurement should be a means to help the sector improve, in that the sector is only as strong as its weakest link. The performance improvements made by nonprofit leaders might generate insights and innovations that will be necessary to improve collective impact. Provided in Table 6 are the nine most important and the nine most unimportant statements for the Weakest Link factor, along with each statement's array position and *z* score. Table G3 with complete statements is located in Appendix G.

Table 6

*Most Important and Most Unimportant Statements and Scores for Weakest Link Condensed*

Item Number	Statement	Array Position	<i>z</i> -Score
39	The sector desperately needs a performance...	5	1.90*
16	Externally driven performance measurement will be...	5	1.81*
18	If NPOs want to prove legitimacy, they need to uphold...	4	1.45
3	A new performance measurement model should help...	4	1.40
21	Measuring and using outcomes data is essential for...	4	1.27
25	NPO leaders must commit to self-discipline and...	3	1.10
7	An effective performance measurement model should...	3	1.09
24	NPO capacity, or lack thereof, impacts effective...	3	1.08
35	Rather than asking traditional strategic planning...	3	1.01
29	Outcome assessment is a highly complicated...	-3	-0.85*
27	NPO performance measurement is receiving increased...	-3	-0.90*
4	A performance measurement system that simplifies the...	-3	-1.07
26	NPO leaders see few frameworks that seem relevant to...	-3	-1.26*
36	Roughly 3,000 charities garner 60% of all revenue that...	-4	-1.31
37	Standardized metrics and reporting systems are...	-4	-1.47*
32	Performance measurement methods are forced upon...	-4	-1.66
41	Too many NPOs are concerned with sustaining...	-5	-1.77*
17	Financial efficiency will always be the most important...	-5	-1.87

*Note:* Numbers with asterisks denote distinguishing statements

Of the four participants comprising Weakest Link, three were women and one was a man. One held a bachelor's degree, and three had master's degrees. All four had 15 or more years of experience in the sector, with one having 20-24 years' and two having more than 25 years' experience. Two of the four participants were executive directors or CEOs at their nonprofit organization, one held a nonprofit senior management position, and one was a foundation professional. Weakest Link represented 13% of the variance.

Members of Weakest Link believed the primary purpose of performance measurement was to help the entire nonprofit sector improve. Organizational learning can help the entire sector grow and improve, wherein strides within one organization can help other nonprofit organizations adopt and implement insights. For Weakest Link, performance measurement is a tool to improve service delivery, resource allocation, and organizational learning (statements 39: +5; 21: +4; 35: +3). The sector needs a model to implement and inform the practices of all nonprofit organizations, and the sector is only as strong as its weakest link (Statements 39: +5, 36: -4; 32: -1). Performance evaluation is valued as a way to strengthen and improve the sector as a whole, where each organization can learn from the insights of others (Statements 18: +4; 31: -1; 36: -4). The sector's weakest links are those organizations that do not engage in performance measurement.

The secondary purpose of performance measurement is for holding organizations and their leaders accountable to their many constituencies. Effective performance measurement should balance the needs of all stakeholders, recognizing the needs and priorities of both upward and downward constituencies (Statements 16: +5; 18: +4; 26: -3). These nonprofit leaders value and recognize their accountability to both donors and



beneficiaries and desire an evaluation practice that does not compromise the needs of any stakeholder (Statements 18: +4; 1: +2, 30: +2; 2: +2). Like the members of Management Tool, the members of Weakest Link were aware of the resource dependencies and sector expectations; however, resource dependencies do not take precedence over service delivery and community needs. These competing needs should be viewed globally in order to maximize value for all concerned parties (Statement 16: +5; 2: +2; 40: 0).

For Weakest Link participants, performance management provides opportunity to achieve the impact donors seek. A new performance measurement model should be standardized. Unlike Road Map and Management Tool, members of Weakest Link do not reject the standardization of performance metrics. Members of Weakest Link dismissed the notion that performance measurement is too complex to standardize across the sector's differences (Statements 37: -4; 26: -3; 8: -2).

In agreement with the other factors, this perspective believes that financial efficiency is not the most important performance metric (Statement 17: -5). These leaders envision a standardized metric that will inform organizational learning and that will be used for reporting performance data to stakeholders (statements 37: -4; 39: +5; 26: -3). Such a model is necessary to hold all nonprofit leaders accountable for measuring performance (Statement 36: -4; 25: +3). From this perspective, metrics will not cause failure, but likely poor commitment by leadership to ensure performance measurement is complete. Buy-in from nonprofit leaders and employees will be necessary for success (Statement 25: +3).

This factor places responsibility on the nonprofit leader to achieve the organization's mission through self-discipline, self-accountability, and creating a

measurement culture (Statement 25: +3). Monitoring and improving performance should be one of the top priorities for organizational leaders (Statement 25: +3). Evaluation should not be viewed as a result of donor trust issues or the desire to hold nonprofit leaders accountable, but rather it is an important responsibility that should not only be expected but also practiced regularly by nonprofit organizations (Statement 32: -4; 41: -5). Further, management structures and systems should be included in a new performance measurement model (Statement 41: -5). These leaders even give consideration to the idea of external evaluation conducted by an unbiased third party in the form of unannounced performance audits (Statement 23: 0). Those sharing viewpoint likely see this as a necessary practice to compel all nonprofit organizations to measure performance. If nonprofit leaders cannot or will not manage and measure their organization's performance, then the opinion of Weakest Link participants would be for these leaders to relegate their positions to make way for leaders who are up to the challenge.

Members of this viewpoint expressed frustration with the continued dialogue and lack of collective action within the sector. These nonprofit leaders expressed that the sector knows enough to make performance measurement a priority. Performance measurement will provide actionable information as well as a path toward mission attainment (Statements 37: -4, 26: -3; 4: -3; 21: +4). Nonprofit leaders should focus on organizational goals and measure organizational outcomes (Statement 21: +4). Those organizations lacking capacity to conduct performance measurement in this manner should be given the necessary support to build organizational capacity (Statement 24: +3). Performance measurement should be both formative and summative to meet the

needs of all constituencies, but best practice would be ongoing performance monitoring by organizational leaders (Statement 38: -1).

An element of frustration is demonstrated because there are still nonprofit organizations that do not engage in the performance measurement practices they know they should use. Members of this factor perceive sector resistance to adopt beneficial performance management behaviors. Weakest Link primarily focuses on why and how to measure performance, however less focus is placed on what the performance metrics should measure. This perspective wants standardization yet fails to focus on specific performance measures to standardize. Weakest Link disagrees that a performance measurement model should not be simplified (Statement: 4: -3). Rather, these participants want a few indicators to provide a sense of organizational outcomes, specifically indicators for organizational health, program performance, and social and economic performance (Statements 7: +3; 5: 1).

Weakest Link members believed nonprofit leaders already know why performance measurement is important to conduct and focus should change to what is measured and how. During a follow up interview, the participant with the defining sort expressed described the ongoing attention to purpose by the sector as “belly button gazing” and further expressed his exhaustion with the question. For Weakest Link participants, purpose is deeply important but continued focus on this element stalls sector progress. Nonprofit leaders need to shift attention to the process and content of performance measurement. Participants of Weakest Link were ready for the sector to accept accountability for performance and to find a way to measure and communicate impact. Standardization of metrics will provide nonprofit organizations a benchmark for

performance, provide a meaningful way to communicate performance data to constituencies, and allow the entire sector to strengthen and grow as a result of meaningful performance insights.

Weakest Link participants believed nonprofit leaders should be able to measure and communicate their performance in terms of the outcomes achieved (Statement 21: +4). Outcome measurement has ties to organizational learning, is easier to implement than most nonprofit leaders believe, and is possible for all nonprofit organizations to achieve. Although achievable outcomes are specific to the organization measured, the standardized expectation would be for leaders of all nonprofit organizations to perform outcome measurement.

**Research Question 2.** The second research question was as follows: What are the characteristics, behaviors, or practices nonprofit leaders believe are critical to developing an effective nonprofit organizational performance measurement model? Although each viewpoint had distinct differences in opinion, the three viewpoints reached a consensus for a few key ideas. Consensus statements represent the Q sort statements with which all three factors either agreed or disagreed. However, the source of agreement or disagreement may be different for each factor. “Different factors may related [*sic*] to different aspects of meaning in the same statement” (Webler et al., 2009, p. 34), and researchers must examine statements within the context of the factor and not just as isolated statements. Consensus occurred among the factors on 12 statements; those most important and most unimportant to effective performance measurement are in Table 7. The themes identified in the results were the overwhelming support for performance management, outcomes measurement, and the need for funding for

performance measurement. Further exploration of these themes will occur in detail in the following section. This section contains a discussion of two statements from the most important and two statements from most unimportant, as these were the statements with an average z score above 1.0 or below -1.0.

Table 7

*Most Important & Most Unimportant Consensus Statements With Array Positions*

Item no.	Statement	Factor 1	Factor 2	Factor 3
3	A new performance measurement model should help the NPO learn and improve performance – to drive internal reflection, improve mission execution, communicate strengths and weaknesses, and enhance adaptability.	5	4	4
7	An effective performance measurement model should include indicators not only for organizational health and program performance, but also for social and economic performance that give a real sense of the organization’s outcomes and progress in meeting its vision.	4	3	3
9	An effective performance measurement model will select measurement indicators tied to the organization’s strategic plan and internal goals.	3	2	1
13	Devoid in the performance measurement discussion is who will pay for it. Since performance measurement is often an underfunded overhead expense, donors who want performance data should match their expectations with their resource commitments so that NPOs can direct resources toward the mission.	1	3	2
24	NPO capacity, or lack thereof, impacts effective performance measurement. A primary objective should be to enhance capacity first so that NPO leaders have the tools to measure performance.	3	1	3
32	Performance measurement methods are forced upon NPO leaders because donors do not trust NPO leaders to manage donations. Increased funding levels raise the burden of proof, forcing leaders to be accountable to those they serve.	-4	-4	-4
36	Roughly 3,000 charities garner 60% of all revenue that comes into the sector each year, and 20,000 garner 85%. The initial goal should be to rate the 20,000 commanding the lion’s share of financial resources and not burden the smaller NPOs with ill-fitted models incompatible with their current capacity.	-3	-2	-4
38	The focus of annual performance measurement is accountability, whereas multiple times per year is focused on performance improvement. While a blend of formative and summative assessment is ideal – how can we expect NPOs to achieve their mission while simultaneously conducting frequent performance measurement?	-2	-1	-1

Members of the three viewpoints agreed that the purpose of performance measurement is for performance improvement and organizational learning. Members of

the Road Map viewpoint indicated performance measurement should help organizational leaders learn how to improve for the sole purpose of getting better at serving beneficiaries. In this viewpoint, performance measurement is a formative self-check for internal purposes. Members of the Management Tool perspective viewed performance measurement as a way to help the organization improve to communicate impact to constituencies. Performance measurement, from this vantage point, is a check that occurs because of resource dependencies. Members of the Weakest Link viewpoint indicated performance measurement and improvement should occur to improve the sector as a whole. Standardization of performance measurement will help strengthen and improve the nonprofit sector, as the application of performance data and improvements can occur across the sector, as evidenced by all three factors positively ranking Statement 3 (ranked +5, +4, +4) as most important to effective organizational performance measurement.

The participants agreed that effective performance measurement should give a real sense of an organization's outcomes and its progress at meeting its vision. According to Road Map participants, these measures were specific to each organization and its leaders' ability to achieve its mission. The measures should have an internal driver and a focus on improving service delivery. Conversely, the Management Tool participants accepted externally driven measures of impact and outcome achievement to support the communication of performance data to funders and other influential stakeholders. The preference for Weakest Link participants was standardized metrics for measuring effectiveness, outcomes, and performance management so that performance data were applicable sector wide. The individuals on all factors positively ranked Statement 7 (ranked +4, +3, +3) as most important.

According to each factor, performance measurement should be a priority for all organizations and not just larger organizations and sector leaders. Members of Road Map indicated leaders in all organizations should measure performance to continue fulfilling community needs. Management Tool members supported leaders in all organizations generating performance data to communicate outcomes to the multiple constituencies who support the organization. Weakest Link members indicated all organizations should have a focus on outcome because the entire sector can benefit from the realizations each organization makes through performance measurement. The three factors disagreed with Statement 36 (ranked -3, -2, -4), which proposed requiring the top 20% of the nonprofit sector to measure their performance.

The individuals on the three factors did not assess accountability as the most important purpose of performance measurement. Members of Road Map indicated the purpose of conducting performance measurement is ultimately for the benefit of the organization's beneficiaries. Conversely, Management Tool participants supported the use of performance measurement as a means for proving the organization's legitimacy to multiple constituencies, most importantly funding agencies. Lastly, Weakest Link members supported conducting performance measurement to prove an organization's legitimacy to the nonprofit sector. The individuals on all three factors demonstrated recognition that real-world performance measurement inherently creates opportunity for stakeholders to react to negative performance, but they show preference for support and guidance to make improvements rather than the imposition of punitive responses. Statement 32 (ranked -4, -4, -4) created a strong disagreement for the three factors. This

statement indicated the purpose of performance measurement is to hold nonprofit leaders accountable.

One statement that did not appear on the PQMethod consensus statement report, but that the three factors strongly reacted to, was Statement 17 (ranked -5, -3, -5). According to Statement 17, financial efficiency is the most important performance metric. Members of each of the perspectives rejected financial performance as the best way to evaluate nonprofit effectiveness. However, none of the participants in the three factors completely dismissed financial efficiency. These viewpoints indicated an understanding that financial performance is necessary to gain a complete understanding of an organization's viability and health.

The three-factor solution generated distinct viewpoints titled Road Map, Management Tool, and Weakest Link. To understand what the participants believed were the most important factors to effective performance measurement, an exploration of consensus statements ensued. First, the viewpoints indicated performance improvement was the primary purpose for performance measurement. Second, performance measurement should be specific to the organization measured. Third, leaders of all nonprofit organizations, regardless of size or status, should conduct performance measurement. Last, accountability was not the most important reason to conduct performance measurement. None of the viewpoints ranked financial efficiency as a critical performance metric.

### **Summary**

Chapter 4 contained the results of the data collection and analysis. Q methodology includes the use of statistical procedures to help uncover the social



perspectives that exist about the topic under investigation. Data analysis ensued using PQMethod 2.33 software (Schmlock, 2002), a review of questionnaire responses, and descriptive statistics. Although commonalities exist between the factors, there is enough difference to represent the individual views of effective nonprofit organizational performance measurement.

The patterns identified through inverted factor analysis were the basis for composing social perspectives (Sickler et al., 2006, p. 356). Q method analysis uncovered a three-factor solution, with the resulting factors named Road Map, Management Tool, and Weakest Link. The chapter included descriptions of the participants aligned with each factor in relation to the study's two research questions identifying what nonprofit leaders believe about effective performance measurement in addition to the behaviors, characteristics, and practices necessary for successful performance measurement. The factors included unique characteristics that supported the theoretical framework and the themes discovered during the literature review.

Eight sorts defined Factor 1, named Road Map because the individuals on this factor viewed performance measurement as a means to help the organization improve. Results from performance measurement provide important information and serve as a formative tool and road map to guide improvements in serving beneficiaries and achieving mission-based goals. Four sorts defined Factor 2, named Management Tool because participants on this factor defined performance measurement as a tool that helps communicate the organization's results to external parties. Participants sharing this viewpoint noted organizational leaders should create performance measurement with the end-state in mind by identifying the mission-driven intended impact they seek to achieve.

Performance management should back performance measurement. Four sorts defined Factor 3, named Weakest Link because the individuals on this factor noted performance measurement should be a means to help the sector improve, in that the sector is only as strong as its weakest link.

The first research question was as follows: What do nonprofit leaders believe to be the most important factors of effective nonprofit organizational performance measurement? Members of each factor had differing opinions about the elements of elements of performance measurement; however, there were some common themes throughout the viewpoints. First, all perspectives focused on performance management. Consistent with all three viewpoints, performance measurement cannot stand alone to be effective. Performance measurement is a tool in the performance management system that should help nonprofit leaders learn and improve organizational performance. The second theme was the unfunded mandate for performance evaluation. The members with shared viewpoints conveyed a desire for a sense of partnership with their funders, both working in concert toward the goal of better service delivery for beneficiaries. Characteristics of an effective performance measurement model are partnership and shared goals. The final theme was outcome measurement. The individuals on each factor favored outcome measurement as the most effective practice for evaluating a nonprofit's performance. When leaders of nonprofit organizations have a better understanding of their outcomes and performance data, then an opportunity exists for greater collaboration and learning, thereby increasing the sector's collective impact.

The second research question was as follows: What are the characteristics, behaviors, or practices nonprofit leaders believe are critical to developing an effective

nonprofit organizational performance measurement model? Although members of each viewpoint had distinct differences in opinion, the participants in each of the three viewpoints reached a consensus for a few key ideas. Consensus statements represent the Q sort statements with which members of all three factors either agreed or disagreed. To understand what the participants believed were the most important factors to effective performance measurement, it was necessary to explore consensus statements. Members of the three viewpoints reached consensus for some key ideas, four of which appeared in detail. First, the members of the viewpoints rated performance improvement as the primary purpose for performance measurement. Second, performance measurement should be specific to the organization measured. Third, leaders of all nonprofit organizations, regardless of size or status, should conduct performance measurement. Last, accountability was not the most important reason to conduct performance measurement. None of the viewpoints ranked financial efficiency as a critical performance metric.

Chapter 5 contains a discussion of the limitations, results, and conclusions of the study. The chapter also includes recommendations for nonprofit leaders and funding agencies. The chapter concludes with a discussion of implications for future research.

## Chapter 5

### Conclusion and Recommendations

Leaders in the nonprofit sector do not agree on how to evaluate nonprofit organizational performance most effectively or on how to define organizational effectiveness (Lecy et al., n.d.). Nonprofit organizations are subject to many stakeholders who often have conflicting expectations (Moore, 2003). These stakeholders include upstream customers, the donors who make NPO work possible, and downstream customers, the clients and recipients of the social benefit produced by nonprofit organizations. Nonprofit organizational evaluation (a) serves to appease funders for their informational purposes rather than for nonprofit organization's performance improvement; (b) is defined and framed differently depending on the source, which creates an inconsistent definition of organizational effectiveness; and (c) is rarely used to change behavior or inspire organizational learning (Ebrahim, 2005). The greatest challenge is finding a balance among upward accountability to donors, meeting the needs of employees, and making positive change in organizational practices (Ebrahim, 2005).

Many concepts exist for how to define and measure nonprofit organizational performance and effectiveness. However, disagreement about to whom nonprofit organizations should be accountable, for what reasons, and how to evaluate nonprofit organizations has resulted in many conceptual and theoretical frameworks for NPO evaluation but little agreement among them (Benjamin, 2008). The centrality of organizational effectiveness in nonprofit organizational performance is generally agreed upon; however, little agreement exists on the definition, measurement, and constitution of effectiveness (Lecy et al., n.d.). Each new construction of nonprofit performance affects

how funders provide resources and contrive expectations, which leaves nonprofit organizations and their leaders in a continual tug-of-war between whose definition and method to implement and what is the best path toward productive evaluation. Although discourse regarding the need for changes and proposed replacement models is ample, little consensus exists within the field on why, what, and how to evaluate while maximizing nonprofit organizational performance, outcomes, and mission achievement.

The general problem guiding this study was that many theoretical models exist that offer different perspectives on nonprofit organizational performance measurement; however, few have been tested or agreed upon or have received notable recognition or adoption within the sector. The specific problem is that a lack of understanding of the subjective opinions held by nonprofit leaders regarding nonprofit organizational performance measurement creates difficulty in designing a widely applicable model that addresses differences in opinion. This lack of understanding might be the reason for the development of multiple unimplemented models. Those leading the shift in these performance expectations and measurement systems must identify the difference in opinions before achieving any consensus or compromise in the creation of a new performance measurement model to guide and assess the sector's performance.

The purpose of the Q methodology study was to identify and collect the subjective perspectives of nonprofit leaders in the United States concerning effective organizational performance measurement practices. Q methodology involves systematic analysis of human subjectivity about a topic of study (McKeown & Thomas, 2013). The intent was to identify patterns of beliefs among nonprofit leaders. A chain referral sample of nonprofit leaders who have experience in nonprofit management and

performance measurement helped to clarify the vast amount of information related to nonprofit organizational performance measurement. From their subjective perspectives, participants identified the most important purpose, content, and process elements of an effective nonprofit organizational performance measurement model.

The study involved collecting the beliefs held by 22 nonprofit leaders, targeting those with experience in nonprofit management and performance measurement within the United States. The research questions for the Q methodology study were as follows:

RQ1: What do nonprofit leaders believe to be the most important elements of effective nonprofit organizational performance measurement?

RQ2: What are the characteristics, behaviors, or practices nonprofit leaders believe are critical to developing an effective nonprofit organizational performance measurement model?

Q methodology includes three main sets of statistical procedures: correlation, factor analysis, and factor score computation (McKeown & Thomas, 2013). Participants bestowed subjective meaning to the Q-set, which was a sample of statements about performance measurement correlated between participant profiles for similar viewpoints (van Exel & de Graaf, 2005). Analysis uncovered patterns from individual responses composed into social perspectives (Sickler et al., 2006). Q sorts represent modeled viewpoints from which data analysis ensued. Chapter 5 includes conclusions based upon the summary of results, limitations of the study, recommendations, significance of the research to leadership, and future implications concerning theory, practice, and research.

## **Summary of the Study**

A sample of 22 nonprofit leaders in the United States who have experience in nonprofit management and performance measurement comprised the sample of this Q methodology study. For the purposes of the study, nonprofit leaders were (a) academic experts who had published within the field in the past 5 years or (b) professionals with 5 or more years of leadership experience within the field. Professionals included nonprofit senior managers, evaluators, grantors and funders, consultants, and watchdog agency leaders.

Participants read and rank ordered 42 Q sample statements from those statements they believed were most unimportant to those they believed were most important to effective nonprofit performance measurement. After participants completed the Q sort exercise, they completed the postsort questionnaire, which included questions about the exercise in addition to demographic information. Data analysis involved PQMethod 2.33 software (Schmlock, 2002), a review of questionnaire responses, and descriptive statistics. Follow-up interviews took place to gain clarity of the viewpoints.

Several test rotations revealed a three-factor solution was the best solution. Sixteen of the 22 sorts loaded significantly on only one of the three factors. Factor 1 had eight significant sorts, whereas Factors 2 and 3 both had four significant sorts. Four sorts were confounded and two sorts were not significant. Researchers consider very high loading on only one factor pure, but consider loading on more than one factor confounded, which represents a hybrid perspective (Giannoulis et al., 2010). The study did not involve using confounded sorts as defining sorts for the factors. Nonsignificant sorts do not achieve a significant load and are those that are not similar to any of the

resulting perspectives (Webler et al., 2009). More details about the data analysis appeared in Chapter 4.

Eight sorts defined Factor 1, named Road Map. The individuals on this factor described performance measurement as a means to improve the organization and stimulate learning. Performance results provide important information, serving as a formative tool and road map to guide improvement. For the Road Map participants, the desire is to better serve their beneficiaries and achieve mission-based goals. Four sorts defined Factor 2, named Management Tool. This factor defined performance measurement as a tool that helps communicate the organization's results to stakeholders. For Management Tool participants, resource dependencies and institutionalized expectations greatly influence their opinion and understanding of performance measurement. Four sorts also defined Factor 3, named Weakest Link. The individuals loading on this factor expressed that performance measurement should be a means to help the sector improve, in that the sector is only as strong as its weakest link. The performance improvements made by nonprofit leaders will generate insights and innovations that will be necessary to improve collective impact.

All three viewpoints, Road Map, Management Tool, and Weakest Link, reflected the belief that the purpose of performance measurement is for performance improvement and organizational learning. Members of the Road Map perspective believed that performance measurement should help the organization learn how to improve for the sole purpose of getting better at serving beneficiaries. In this viewpoint, performance measurement is a formative self-check for internal purposes. Members of the Management Tool perspective viewed performance measurement as a way to help the



organization improve for communicating impact to constituencies. Performance measurement, from this vantage point, is a check that occurs because of resource dependencies to fulfill donor expectations for performance data reporting. Members of the Weakest Link perspective believed performance measurement and improvement should occur to improve the sector as a whole. Standardization of performance measurement will help the sector to learn and grow, as performance data and improvements can apply across the sector.

Participants from all three perspectives agreed effective performance measurement should give a real sense of an organization's outcomes and its progress at meeting its vision. Members of the Road Map perspective defined these measures as specific to each organization and its ability to achieve its mission. Additionally, the measures should be internally driven and focused on improving service delivery. Conversely, participants who shared the Management Tool viewpoint accepted externally driven measures of impact and outcome achievement to communicate performance data to funders and other influential stakeholders. Members of the Weakest Link perspective preferred standardized metrics for measuring effectiveness, outcomes, and performance management so that performance data were applicable sector wide.

The individuals who shared each factor believed performance measurement should be a priority for all organizations and not just larger organizations and sector leaders. In accordance with the Road Map viewpoint, leaders of all organizations should measure performance to continue fulfilling community needs. Participants who shared the Management Tool perspective thought all organizations should generate performance data to communicate outcomes to the multiple constituencies who supported the

organization. Participants who shared the Weakest Link viewpoint reasoned that all organizational leaders should have an outcome focus because the entire sector can benefit from the realizations each organization makes through performance measurement.

The three factors do not reflect a belief that accountability is the most important purpose of performance measurement. The Road Map participants' viewpoint was that the purpose of conducting performance measurement is ultimately for the benefit of an organization's beneficiaries. Conversely, for Management Tool participants, performance measurement data are a means for proving an organization's legitimacy to multiple constituencies, most importantly funding agencies. Lastly, the Weakest Link participants' perspective included the belief that performance measurement helps to prove an organization's legitimacy to the nonprofit sector. Participants from all three perspectives understood that in the real world, performance measurement inherently creates opportunity for stakeholders to react to negative performance, but they desire support and guidance to make improvements rather than to impose punitive responses.

The three viewpoints emanating from this study reflected nonprofit leaders' acknowledgment that their organizations are part of an open system, and their actions are part of a larger system or network of actors. Further, the viewpoints identified issues and gaps within the system with the specific goals to address these social issues. Through their viewpoints, the participants demonstrated a desire for new standards and expectations for their nonprofit organization's performance and noted the current institutionalized standards and expectations do not serve the best interest of mission effectiveness. Finally, through the viewpoints, participants revealed the upward and downward resource dependencies nonprofit organizations are accountable to, the desire

for increased collaboration, and the need for a new construction of nonprofit organizational effectiveness and performance expectations.

Each of the perspectives rejected financial performance as the best way to evaluate nonprofit effectiveness. However, none of the participants in the three factors completely dismissed financial efficiency. The participants understood financial performance is necessary to gain a complete understanding of an organization's viability and health.

### **Discussion of Results**

Discovered in this study on nonprofit leader opinions was the acknowledgment of a broken system of nonprofit organizational performance expectations and performance measurement. Insights from the nonprofit leaders revealed the characteristics of a new system for generating meaningful nonprofit performance data. The system envisioned in the viewpoints could strengthen performance management through outcome measurement. The participants believe such changes could promote organizational learning, inspire collaboration and partnership with funders, while including beneficiaries and their feedback. Through each viewpoint, participants exhibited different opinions about the elements of performance measurement, but there were some common themes throughout the viewpoints. Three themes identified in the viewpoints were the need for performance management systems, the unfunded mandate for performance measurement, and the desire to use outcome measurement to measure nonprofit effectiveness.

Nonprofit organizations are social systems dependent upon external resources for survival (Scott & Davis, 2007). External influence and pressure threaten autonomy and demand conformity to and institutionalization of external standards and norms (Worth,

2009). The standards and definitions of effectiveness, impact, and performance are dynamic, subjective social constructions (Herman & Renz, 1997) that help to explain the ongoing dialogue regarding how to evaluate nonprofit organizational performance most effectively. The viewpoints uncovered in Chapter 4 were reflective of the open systems these nonprofit leaders operated within and demonstrated awareness of the effect of their socially created and institutionalized resource dependencies. The three viewpoints identified in this study demonstrated an understanding that their nonprofit organizations are part of a larger network in which not all forces are under their control.

Resource dependence theory helps explain the external demand for accountability and performance measurement in the nonprofit sector (Worth, 2009). Most nonprofit leaders desire independence for their organization to tackle the problems their mission was created to tackle, but this becomes hindered by their resource dependencies (Pratt, 2007). Institutional theory accounts for nonprofit leaders adopting external standards and expectations and the tendency for organizational leaders to assimilate and adhere to external standards and norms (Worth, 2009). Dependence creates the need to demonstrate responsiveness and incentive to devote more time and attention to some stakeholders rather than to others, especially when catering to an external governance agenda that often consumes a disproportionate amount of time and resources (LeRoux, 2009). The Road Map, Management Tool, and Weakest Link perspectives were mindful of their resource dependencies and expressed the desire and need for a balanced partnership with funders. An ideal partnership would help fund evaluation expenses and enable nonprofit leaders to build performance management systems.

The three perspectives focused on performance management as the complete solution, with performance measurement as a singular part of increasing organizational effectiveness. The Road Map, Management Tool, and Weakest Link viewpoints demonstrated that performance measurement cannot stand alone but rather must be a part of a performance management system. A performance management system should help organizational leaders learn and improve performance. Performance management is the larger system for generating effective performance, with performance measurement being one of the system's practices.

**Performance measurement for Road Map Leaders.** The Road Map perspective defines effective performance measurement as a mission-based, community-focused, and formative evaluation model. This model would be organization specific and inform performance improvements specifically targeted for goal attainment. For members of Road Map, the primary focus is meeting community needs. Inclusion of feedback from downward constituencies is critical to this process. While this model is largely an internal exercise conducted by the organization's members, these leaders welcome external input to prevent bias from skewing results and maintaining status quo. Ideally performance measurement data would generate organizational learning, innovation, and adaption to meet ever-changing community needs.

To implement such a model, numerous considerations would need to be made. First, organizations adopting this model would need to select funding partners with closely aligned goals. Although nonprofit leaders will never eliminate resource dependencies, they can minimize the power dynamic by selecting funders with truly complementary goals. This means nonprofit leaders will have to decline funding from

some donors, including funders who pull the organization away from the mission-based goals.

Funder relationships should be a partnership with complementary rather than competing goals. Funding must be available as part of existing revenue streams and with all new monies so that nonprofit organizations can build performance management systems and improve performance evaluation. Ideal partners will view administrative costs as necessary to produce good outcomes. Partnership will be critical for shifting performance and funding expectations. Nonprofit leaders will need to educate their communities about the true costs associated with producing impact and help them to understand how damaging the focus on overhead is for their organizations.

Further, nonprofit leaders would need to develop a performance culture within their organizations. Cultural changes that are driving mission effectiveness will not be enough to change the internal practices of individual nonprofit organizations. Nonprofit leaders will need to embrace the exciting and humbling experience of transforming their organizations from program-based thinking to a social change orientation (Wellar, 2012). Some nonprofit leaders will resist making this change, which will threaten their organizations' relevance and continued existence. Rather than fighting to maintain the status quo, nonprofit leaders should embrace and lead the sector's transformation. Funders should help incentivize the change rather than discourage organizational leaders from avoiding the risk.

**Performance measurement for Management Tool leaders.** The Management Tool perspective defines effective performance measurement as a performance management tool that enables evidence-based decision-making. Such a model would

focus on the desired end-state, promote performance improvement, and help the organization attain mission-driven intended impact. Strategic goals should be created in concert with upward constituencies, namely prominent funders. This relationship would be collaborative, with funding agencies serving as mentors and partners for goal attainment. Performance data would be used to communicate the organization's social impact and would help to prove the organization's legitimacy to stakeholders. Performance metrics would be organization specific and outcome measurement would be the primary method for collecting and communicating performance data.

Hamel and Prahalad (1989) described strategic intent as the objectives and behaviors an organization will define and execute to become an industry leader, including the active management process that defines success, motivates employees through shared vision, allows for innovation, and evolves over time. Nonprofit leaders need to create a strategic clarity that drives every facet of the organization, from marketing to accounting, from resource expenditures to staffing decisions, and so forth, which often might involve making substantial changes, including ending programs or services or reorganizing the organization to achieve the organization's strategic intent.

To implement the Road Map model, nonprofits would need to partner with funders committed to advancing mission progress. Funders with incongruent or competing goals would prevent this model from succeeding. Consideration would also need to be made for funding partner performance data reporting expectations. Funders should be mindful of the data they require and ensure that metrics do not place focus away from strategic intent. In turn, funding partners would need to make additional funding available to make this process achievable. The primary focus would be to create

a larger performance management structure to help nonprofit leaders inform performance improvement and meet upward constituency performance expectations.

**Performance measurement for Weakest Link leaders.** The Weakest Link perspective defines effective performance measurement as a standardized process to communicate the sector's collective impact as well as individual organization performance, in which all nonprofit organizations would be expected to participate. This model would focus on service delivery improvement, improved resource allocation, and promoting organizational learning. Because of this focus, the model presented by Weakest Link would balance both upward and downward accountabilities. The standardized expectation would be for nonprofit organizations to measure mission-based outcomes as part of a larger performance management structure. Ideally, this model would create comparative data to hold nonprofit leaders accountable for performance monitoring and improvement.

To implement such a model, the expectation would need to be set for all nonprofit organizations to comply with this standardized process. Likely government legislation would be required. Without a mandate to participate requests alone would likely fail to gain sector-wide traction. Where the models proposed by Road Map and Management Tool were organization specific and easily adopted by organization leaders, the model proposed by Weakest Link would require greater dialogue regarding how to implement this model and who would be responsible for regulating and monitoring compliance, defining penalties for noncompliance, and a mechanism for reporting and comparing performance data. An unbiased, nongovernmental, third party such as a charity watchdog agency could be enlisted for this process.



**Commonalities and implementation considerations.** Subjectivity influences the conceptualization of organizational effectiveness, which is not a stable or definite construct as new interactions influence the individual's understanding and definition of the construct. The accepted nonprofit effectiveness construct places ample attention on overhead ratios. Many nonprofit leaders and thought leaders reject overhead, yet this has not lost traction with donors and the media to the extent it has within the sector (Mittendorf, 2013).

The leaders in this study agreed that financial efficiency, also referred to as overhead, is an ineffective metric for mission effectiveness. The leaders attributed most of the necessary management and infrastructure investments that would improve mission effectiveness to overhead. Minimizing overhead creates a deficit in the ability to deliver superior service to constituents (Goggins-Greggory & Howard, 2009). Organizational leaders underspending on overhead or underreporting expenses create unrealistic expectations of the true operational cost of outcomes, which is a harmful practice called the nonprofit starvation cycle that has led to unrealistic funder expectations (Goggins-Greggory & Howard, 2009).

When most funding agendas do not include basic operational and infrastructure funding, it becomes clear why the leaders of most nonprofit organizations minimally or ineffectively practice performance measurement. Performance measurement, let alone creating a performance measurement system, is a luxury that underfunded organizations cannot afford. In June 2013, the presidents of the leading charity watchdog agencies issued a joint letter titled *The Overhead Myth* calling for an end to using overhead ratios as a sole indicator of nonprofit effectiveness (Taylor, Harold, & Berger, 2013). This

letter explained that “...*how* money is spent is often not a very reliable indicator of the outcomes achieved by the nonprofit” (Larkin, 2013, para. 14). This is a good start in the right direction toward effective performance measurement. However, additional donor education and funding, specifically for performance measurement or generally as increased levels of general operating funding, is necessary for the performance measurement models envisioned by the study’s three perspectives to become widely accepted and regularly practiced.

Members of the three viewpoints expressed how funders expect performance data as a condition of the funding, yet few grantors include funding for administrative and evaluation purposes. The participants conveyed a desire for a sense of partnership with their funders, both working in concert to provide better service delivery for beneficiaries. Characteristics of an effective performance measurement model are partnership and shared goals.

Partnership with funding agencies would be an appropriate place to start in creating a mission-based performance management system. Such a system would include upward and downward dependencies in the system’s creation and ongoing use. This system would promote organizational learning through feedback loops and could help to measure and improve organizational outcomes. Without funding partner mentorship and financial support this would be an impossible undertaking for most resource strapped nonprofit organizations. Economic challenges often present an impetus for nonprofit leaders to accept all funding, despite the potential for mission creep or inconvenient expectations and demands. A long-term funding strategy will help leaders determine what type of funding relationships are most suitable for the organization, in addition to

identifying those resource dependencies the organization should avoid. Increased collaboration between funders to determine duplication in funding agendas would help minimize redundant funding, as well as redirect funding to issues that are systematically ignored (Sherry, 2013).

The Road Map, Management Tool, and Weakest Link perspectives desired organizational-specific measures linking concretely to the mission rather than ill-fitting and vague standardized metrics to evaluate nonprofit organizational performance. The factors favored outcome measurement as the most effective practice for evaluating a nonprofit's performance. When leaders of nonprofit organizations have a better understanding of their outcomes and performance data, then there is opportunity for greater collaboration, learning, and increasing the sector's collective impact.

Nonprofit leaders are increasingly ready to make the change to a different focus for performance measurement and reporting. Discovered in the viewpoints identified in this study, nonprofit leaders desire a new definition based upon outcomes and social impact. Members of Road Map, Management Tool, and Weakest Link preferred outcome measurement as a means to assess and improve performance. The participants' viewpoints indicated an eagerness to change the expectations and dynamics to a system that will enable impact rather than stifle it.

Members of Road Map, Management Tool, and Weakest Link agreed that the purpose of performance measurement is performance improvement. For performance improvement to occur, there must be a feedback mechanism that informs organizational learning. In open systems, feedback is critical to sustenance and growth, and measurement is an essential method for feedback collection. However, feedback differs

from measurement in that feedback is life-sustaining, contextual, dynamic, and self-generated (Wheatley & Kellner-Rogers, 1999). Measurement tends to be rigid and static, whereas feedback is dynamic and adaptive, which is more suitable for addressing the challenges of the current climate.

The accepted system is largely activity driven rather than results oriented, and investments have essentially made services available opposed to driving change (Berger, Penna, & Goldberg, 2010). The shift to organization-specific outcome measurement will require nonprofit leaders to define the value of their organization's work and then determine a way to measure the value created (Berger et al., 2010). Still, the proposed shift is untested, and its acceptance and potential for success are unknown (Schambra, 2013b).

The performance measurement models envisioned by Road Map, Management Tool, and Weakest Link leaders would help to differentiate effective nonprofit organizations from their underperforming peers. These models would generate performance data that could help to minimize competition for scarce resources, prevent duplication of services, and lead to an increase in scale for those programs and services that have proven successful. Additionally, outcome data could continually drive performance improvements, drive higher standards for service delivery, and generate greater levels of social impact.

Organizational leaders must be bold enough to assess their competencies and weaknesses honestly, abandon the status quo, maintain focus on strategic intent, eliminate distractions, and diversify revenue streams (Edgington, 2010). Often more uncomfortable, leaders must also demand a working and financially invested board of

directors, fire underperforming staff, be candid about the true cost of outcomes, and commit to boldness and impact rather than mediocrity (Edgington, 2010). As evidenced in the study's viewpoints, nonprofit leaders must focus on outcomes and not activities. Managing performance is a cyclical process that requires definition, measurement, learning, and improvement (Tierney & Steele, 2011). The process will require challenging assumptions, finding ways to collect data from the front lines of the organization and from beneficiaries, and ensuring every member of the organization has the same focus and drive for the strategic intent (O'Donovan & Rimland-Flower, 2013).

The exciting part about the shift toward outcome and impact measurement is that these metrics are the point where quantitative and qualitative converge. Outcome and impact measurement can satisfy all donor types; the segmentation of the donor types and the presentation of the information are important. After creating a mission-based performance culture, implementing a performance management system, and collecting performance data, organizational leaders would be remiss not to spend the extra time to communicate results. Nonprofit leaders will need to create the message of their outcomes and impact within the community. A shift from the overhead definition of effectiveness would better allow nonprofit leaders to spend on marketing and advertising to communicate results.

Morino (2011) explained the shift to outcome measurement will require funders to think big and help drive nonprofit mission effectiveness rather than solely require occasional performance data that do little to help leaders of nonprofit organizations achieve their goals. Funders and watchdog agencies must be advocates for meaningful dialogue, eliminate reporting expectations that generate meaningless data, and ensure

performance management and measurement tools will help leaders of nonprofit organizations create social value (Berger, 2013).

Effective collaboration will require shared responsibility from both nonprofit organizations and funders. Leaders of nonprofit organizations will need the resources to achieve their goals, productive working relationships based upon strategic clarity, and conditions enabling organizational learning and improvement (Tierney & Steele, 2011). The partnership must allow for the learning that occurs from failures so that nonprofit organizations will try different ways to achieve their intended impact.

For many nonprofit organizations the bigger problem is a need to develop a performance management culture and most nonprofit organizations identify leadership development as their most glaring organizational challenge (Bibb-Binder & Kramer, 2013). Investments in infrastructure and leadership development will further drive social impact. Mentorship and collaboration from funders will be vital to facilitating the shift to outcome and impact measurement. Leadership development must be an iterative process with opportunities to plan, learn, measure, and adjust. Developing this monumental change will require more than a single annual gift, but rather a longer term commitment to mission achievement (Bibb-Binder & Kramer, 2013).

### **Significance of the Study for leadership**

Nonprofit leaders must take an active role in creating a culture of performance management that facilitates performance measurement and performance improvements if they are to advance the mission of the organizations they lead. Performance measurement is one of the most critical practices of an organization to determine and expose value creation and destruction, and as such, organizational leaders must make

measurement leadership a fundamental part of their jobs (Spitzer, 2007). Performance management was the top priority for the participants in the study. Above all else, these nonprofit leaders believed performance management is most important to effective performance measurement.

Lack of focus on performance is a sector-wide problem. Failure to manage to outcomes reduces the sector's collective impact (Morino, 2011, p. 1). Focus on performance metrics is important, because what leaders measure is what ends up being managed (Bladt & Filbin, 2013). Assessment will help nonprofit leaders to “identify successes and opportunities for improvement; jump-start a change initiative or energize current initiatives; energize the workforce; focus the organization on common goals; assess [the] organization's performance against the competition; align . . . resources with . . . strategic objectives; [and] deliver world-class results” (National Institute of Standards and Technology, 2010, para. 5).

For nonprofit organizations, this change is unavoidable. Access to information, increased demand for transparency and accountability, and added competition have catapulted nonprofit performance to the forefront. Driving the need for assessment are demands for mission effectiveness and the rise of impact investing, where investors donate based upon rate of return and social impact (Morino, 2011). The challenge is to not be stuck on performance measurement as the end goal, but rather to formulate it as a tool in a greater performance management system that helps nonprofit leaders provide increasingly improved service (Morino, 2011). As identified by the study's participants, financial efficiency will no longer suffice as a metric for organizational effectiveness.

The rise of the digital information age “has left organizational leaders with a real problem, since the trusted, traditional approach to strategic planning is based on assumptions that no longer hold. The static strategic plan is dead” (O’Donovan & Rimland-Flower, 2013, para. 7). Instead, nonprofit leaders should focus on the organization’s strategic intent. Strategic intent begins with the organizational mission and the greater impact the organization seeks to have for beneficiaries. Leaders need to be bold, set ambitious goals based upon the long-term vision, and then work backward to enable the organization to meet that vision (Govindarajan, 2012). Well-developed and poorly implemented strategic plans often remain in a forgotten binder on a dusty shelf. These static plans will no longer be sufficient. Leaders need to drive flexibility and adaptability and incorporate organizational feedback and tools that will achieve each milestone in the path toward mission accomplishment. After the definition of strategic intent is clear, implementing outcome measurement will become a less arduous task.

Superior performance does not happen spontaneously, and, if it occurs by accident, it is hard to replicate. Measuring inputs and outputs helps nonprofit leaders to realize the need to measure outcomes. With practice, active effort, and diligence, nonprofit leaders can effectively implement outcome measurement that will lead to nonprofit organizations being better able to communicate their impact. Performance measurement must be part of a broader performance management system informed by the organization’s strategic intent, the mission, and the social problems the organizational leaders seek to address.

The perspectives collected in this study identified the characteristics of three unique performance measurement models in addition to some common performance



measurement characteristics. These commonly agreed upon characteristics could help to reinvent a performance measurement model. The characteristics of effective performance measurement identified in this study could be the starting point for broader creation and implementation throughout the sector. Further, these models are likely to appeal to the nonprofit leaders responsible for implementation of a new measurement system.

Caution is necessary during the shift toward intended impact and outcome measurement. Many of those engaged in the performance measurement dialogue have considered measuring impact as being elusive or too difficult for most nonprofit organizations to manage. Discussion and experimentation have led to the belief that measuring organizational outcomes offers nonprofit leaders a means to document impact (Miller-McLemore, 2012), which the opinions expressed in this study further supported, demonstrating preference for outcome measurement.

### **Recommendations to Nonprofit Leaders**

Although there were differences in viewpoints identified in this study, the commonly identified elements of effective performance measurement would be a smart place for nonprofit leaders to use as a starting point for leading the change within their organizations. Leaders of an organization must first determine the organization's intended impact. Intended impact defines "what the organization is trying to achieve and will hold itself accountable for within some manageable period of time. It identifies both the *benefits* the organization seeks to provide and the *beneficiaries*" (Colby, Stone & Carttar, 2004, para. 9). Shifting focus to intended impact driven by theories of change is an iterative process that should start with open-ended questions to help organizational leaders determine which activities and beliefs are most important to mission effectiveness

(Colby et al., 2004). Efforts to achieve strategic clarity can be messy and contentious because it is an exercise in changing the organizational culture (Morino, 2011). This process will also help leaders identify their organization's baseline and establish a timeline to achieve milestones (Axelrod, 2013).

Members of Road Map stressed the importance of nonprofit leaders using performance data to guide, strengthen, and improve nonprofit organizations but also showed awareness that not all organizations have the same baseline capacity. For evaluation practices to thrive, organizational leaders must first build the capacity to bring discipline, structure, and process (Kramer & Stid, n.d.). Funders should fund performance management training and infrastructure as their investment in the grantee's mission and subsequent mission effectiveness.

The participants of the Management Tool perspective highlighted the importance of including stakeholders when developing an organization's intended impact. Far too often a critical voice is left out in the assessment process. More often than not, nonprofit leaders do not consult with beneficiaries during the evaluation of the programs and services they receive. This is not a revolutionary concept; leaders in the corporate sector have created, tested, and perfected many systems to collect and analyze consumer feedback effectively (Birdsong & Ni, 2012). Members of Road Map stressed the importance of including beneficiaries in the performance measurement process so that the most critical feedback is collected. Nonprofit leaders learn the most important performance improvements from those beneficiaries who are rarely asked (Crawford, 2013). Incorporating beneficiary feedback can further reduce the imbalanced power dynamic between nonprofit organizations and funders and provide critical insights

grounded in the daily experiences of those the programs exist to serve (Twersky, Buchanan, & Threlfall, 2013).

There are many reasons to omit beneficiaries' voice, including the expense, the difficulty in gathering responses or recruiting participation, and the fear of what they might say (Twersky et al., 2013). No matter the reasons for avoiding beneficiary feedback, within a new performance management system, this needs to be the most sought out source of feedback. Leaders of nonprofit organizations must invest in the tools and systems that will allow beneficiary feedback to be a much easier and effective source of performance data. These leading indicators

are important because they allow decision makers to make improvements while the program is under way, rather than waiting until after negative outcomes . . . to make adjustments. Beneficiary feedback isn't just the right thing to do; it's the smart thing to do. (Twersky et al., 2013, p. 45)

The members of the Weakest Link factor highlighted the shift toward impact measurement and organizational learning as an opportunity to enhance and increase the sector's collective impact capability. Measures of nonprofit performance must account for the social benefit that nonprofit organizations generate collectively (Rosenman, 2013). If the performance management system includes collective impact, nonprofit leaders will be more likely to create a broader culture of learning, share good practices, collaborate with other NPOs, and influence peers to multiply the collective impact (Smith-Milway & Goggins-Gregory, 2013). As organizations within an open system, the future will require networked systems that communicate and work toward greater effectiveness collectively (McCambridge, 2012a).

For a sector driven by inspiration and vision of a better world, passion drives the actions of many nonprofit leaders and their organizations, rather than the analytical mechanisms that drive productivity and effectiveness (Kramer & Stid, n.d.). Focus on outcomes and social impact is where quantitative meets qualitative, where passion intersects with measurement, and would prime the sector for reporting data from which most stakeholders would gladly receive and benefit.

The leaders in this study elevated performance improvement as the ultimate goal for performance measurement. Measurement is not an end in itself, but rather should support an organization's need for quality feedback. Feedback enables organizational learning. Organizational learning prevents inertia and prevents people from getting set in their ways. Trust and a belief that punitive action will not ensue from imperfect results must also exist so that nonprofit leaders feel encouraged to prioritize feedback and share results. When feedback receives the utmost priority in the nonprofit sector, then the leaders will build the appropriate capacities, implement performance management systems, include beneficiaries and elevate their status, measure outcomes and impact, and amplify collective impact.

Becoming an effective organization essentially equates to continuous improvement and outcome achievement. No matter the organization's status, strengths, or weaknesses, a purposeful and holistic effort to manage performance and encourage organizational learning will be worthwhile for employees, volunteers, and beneficiaries alike (Kramer & Stid, n.d.). The three perspectives uncovered in this study provide three descriptions of effective performance measurement models that would be a smart starting

point for nonprofit leaders to shape their organizations' performance management system.

### **Limitations of the Study**

At the time of this study the body of research had gaps in the identification of nonprofit leaders' viewpoints of effective nonprofit organizational performance measurement. Uncovering these perspectives of the topic helped to fill in part of the research gap. There were several limitations within the study, as presented in this section.

In this study, for the purposes of identification and selection, the population of nonprofit leaders met certain criteria, and those leaders who the researcher contacted and who agreed to participate further limited the study. Although the sample was small, was purposive, and might not have been broadly applicable to individuals with different experiences (Akhtar-Danesh et al., 2008), the goal of this Q methodology study was to identify different patterns of thought rather than the numerical distribution within a larger population (Akhtar-Danesh et al., 2009).

Time was another limitation, as those leaders who had time to share at the time of the research participated and those with busier schedules were unable to participate. Those leaders who were busier or unsolicited for participation might have had differing opinions than those who contributed to the results. As it would have been cost and time prohibitive to study the entire population, a limitation of the study population was the inclusion of a portion of the population that was accessible and available to the researcher (Cooper & Schindler, 2008). Further research with different recruitment methods could improve upon this limitation.

A limitation of the study was participant response to the statement deck and the questionnaire. Some participants stated the language used or ideas presented were confusing or unclear. Others expressed that the exercise was more difficult than they had originally anticipated. Further, participants might not have understood the instructions or had minimal knowledge of the topic, despite their position. Participants might have felt a sense of incompetence, experiencing difficulty completing the exercise or representing their viewpoint. The definitions of validity and reliability are different in Q methodology because the expressed perspective derived from the participant's viewpoint rather than the researcher's external standpoint (S. R. Brown, 1980; Hathcoat & Montgomery, 2010). Validity is only an issue when participants genuinely fail to represent their own perspective (Choi, 2010; Hathcoat & Montgomery, 2010).

Another limitation was the design of the study, in that mailing the Q sort materials and relying on e-mail and phone interviews minimized the interaction between the researcher and the participants. Completing the Q sort and conducting the postsort interview in person might have reduced some of the confusion. Conducting the exercise in person would have allowed for a richer dialogue and a richer understanding of the participants' viewpoints. Incorporating face-to-face appointments that dedicated a block of time for the Q sort and interview might have helped to avoid possible distractions that might have occurred during the independent sorts. The degree of honesty of participant responses to the survey was a limitation. Despite reassurances of confidentiality, some participants might have feared being too honest or revealing of their true opinions about performance measurement.

Researcher bias in the selection of Q-statements from the literature review research might have been another limitation of the study. Bias might be a limitation because the bias might reflect a researcher's interests rather than providing a complete sample from the concourse (McKeown & Thomas, 2013). The systematically constructed structured sample of statements explained in greater detail in Chapter 3 reduced or mitigated this bias. These limitations might have prevented the generalization of results, but the findings of the study are still pertinent to understanding nonprofit leaders' perceptions of effective nonprofit organizational performance measurement.

### **Directions for Future Research**

Effectiveness, performance management, and performance measurement are not new topics within the nonprofit sector, but political, cultural, and economic changes are driving these topics to greater prominence. The changes are driving the need for further study of the opinions of nonprofit leaders because they are the change agents within nonprofit organizations. Five possibilities for future research follow.

One option is to replicate and improve this study by clarifying and simplifying the language used in the Q-set. A future researcher could simplify the Q-set and replicate the study. The replicated study might include different viewpoints based upon the change in language as well as the individuals recruited.

A second possibility is that it may be beneficial to determine if viewpoints are contingent upon nonprofit leaders' sector roles. A researcher could replicate the study and only change the target for recruitment. A researcher could also repeat the study and target participants based upon their specific role (e.g., executive director, foundation official, consultant) in the sector rather than include all roles within the participant group.

A third option is to replicate this study and target nonprofit leaders based upon their National Taxonomy of Exempt Entities classifications. Viewpoints may differ drastically based upon the difference in mission-based activities. Creating a standardized performance measurement model may be more possible if based upon National Taxonomy of Exempt Entities classifications.

The fourth possibility for future research would be to determine whether gender plays a significant role in which viewpoint a person identifies with. While the study had a small population sample and is not generalizable, two of the three factors were predominately one gender or the other. Factor 2 was composed of three men and one woman. Conversely, factor 3 was composed of three women and one man. A researcher could design a future study to account for gender differences and identify whether gender plays a role in nonprofit leader opinions about effective performance measurement.

The final possibility for future research is to conduct survey of the sector to determine the distribution of viewpoints among the larger population of nonprofit leaders, which might also help to understand how many organizations are experiencing the various barriers preventing performance management and the implementation of performance measurement. Using the viewpoints discovered in this study, a researcher could survey a statistically representative sample of nonprofit leaders to determine the distribution of these viewpoints among the U.S. nonprofit leader population.

### **Summary of the Study**

Leaders in the nonprofit sector do not agree on how to evaluate nonprofit organizational performance most effectively and how to define organizational effectiveness (Lecy et al., n.d.). The greatest challenge is finding a balance among



upward accountability to donors, meeting the needs of employees, and making positive change in organizational practices (Ebrahim, 2005). Current evaluation and assessment metrics often hurt rather than help leaders of NPOs improve organizational performance (Ebrahim, 2005) and provide external parties with a limited or incomplete understanding of outcome delivery (U.S. General Accounting Office, 2002). A lack of understanding of the subjective opinions held by nonprofit leaders regarding the purpose, process, and content of nonprofit organizational performance measurement creates difficulty in designing a widely applicable model that addresses differences in opinion.

The purpose of the Q methodology study was to identify the subjective perspectives of nonprofit leaders in the United States concerning effective organizational performance measurement practices. Q methodology involves systematic analysis of human subjectivity about a topic of study (McKeown & Thomas, 2013). The intent was to identify patterns of beliefs among nonprofit leaders. The study involved collecting the perspectives held by 22 nonprofit leaders and targeting those with experience in nonprofit management and performance measurement within the United States.

Q methodology includes three main sets of statistical procedures: correlation, factor analysis, and factor score computation (McKeown & Thomas, 2013). Participants bestowed subjective meaning to the Q-set, a sample of statements about performance measurement, which the researcher correlated between participant profiles for similar viewpoints (van Exel & de Graaf, 2005). Analysis uncovered patterns from individual responses subsequently composed into social perspectives (Sickler et al., 2006).

Several test rotations indicated a three-factor solution was the best solution. Sixteen of the 22 sorts loaded significantly on only one of the three factors. Factor 1 had

eight significant sorts, whereas Factors 2 and 3 both had four significant sorts. Four sorts were confounded, and two sorts were nonsignificant.

Eight sorts defined Factor 1, named Road Map. The individuals on this factor described performance measurement as a means to improve the organization and stimulate learning. Performance results provide important information, serving as a formative tool and road map to guide improvement. For the Road Map participants, the desire is to better serve their beneficiaries and achieve mission-based goals. Four sorts defined Factor 2, named Management Tool. This factor defined performance measurement as a tool that helps communicate the organization's results to stakeholders. For Management Tool participants, resource dependencies and institutionalized expectations greatly influence their opinion and understanding of performance measurement. Four sorts also defined Factor 3, named Weakest Link. The individuals loading on this factor expressed that performance measurement should be a means to help the sector improve, in that the sector is only as strong as its weakest link. The performance improvements made by nonprofit leaders will generate insights and innovations that will be necessary to improve collective impact.

The three themes found in the viewpoints were performance management, the unfunded mandate for performance evaluation, and outcome measurement. For performance measurement to work, it has to be part of a broader performance management culture and system. Nonprofit leaders desire a collaborative partnership with funders, and funding is necessary for effectively performance measurement to occur. A focus on outcome measurement will allow leaders of nonprofit organizations to communicate their impact as well as increase the sector's collective impact.

Because they are open systems, nonprofit organizations are subject to a vast pool of interested stakeholders, each with their own opinions and beliefs about nonprofit organizational performance and effectiveness. Resource dependencies affect which opinions and beliefs nonprofit leaders hear most and which socially constructed performance expectations and standards nonprofit organizations institutionalize. To truly improve nonprofit organizational performance, performance measurement must be part of a larger performance management system. Performance measurement needs financial support to become common practice throughout the sector. Outcome measurement connected to the organization's mission-based goals is the best measure of nonprofit organizational effectiveness.

### **Concluding Remarks**

As history has shown, paradigm shifts take time to reach critical mass. Decades of searching for a definition for nonprofit effectiveness have led many nonprofit leaders to believe that social impact is the answer. Although the paradigm of nonprofit effectiveness has slowly evolved over the past few decades, technological and socioeconomic forces, in addition to the overhead myth, are necessitating the change take shape much quicker (Taylor et al., 2013).

The sector would benefit from a shift in the dialogue between nonprofit organizations and donors from how much the leaders spend on overhead to a discussion about what good outcomes cost (Goggins-Greggory, 2013). The need to make this change is not new, and has been a topic of discussion for more than a century, dating back to

Carnegie's complaint that the vast majority of giving is wasted, and to Rockefeller's insistence that we stop just treating the symptoms of problems and instead get to their root causes so that we could solve them once and for all. (Schambra, 2013a, para. 31)

The situation has never been more primed with the tools, technology, and momentum to make the transition smoother.

Discovered in this Q methodology study of nonprofit leader opinions was the acknowledgment of a broken system of nonprofit organizational performance expectations and performance measurement. Insights from the nonprofit leaders demonstrated characteristics of a new system for generating mission-driven impact. The system envisioned in the viewpoints could strengthen performance management, promote organizational learning, inspire collaboration and partnership with funders, include beneficiaries and their feedback, and focus on outcome measurement. Each of the factors exhibited differing opinions about the characteristics, behaviors, and practices; however, there were some common themes throughout the viewpoints. Three themes identified in the viewpoints include the need for performance management systems, the unfunded mandate for performance measurement, and the desire to use outcome measurement to measure nonprofit effectiveness.

First, the focus of all three perspectives identified in the study was performance management as the complete solution, with performance measurement as a singular part of increasing organizational effectiveness. The viewpoints demonstrated performance measurement cannot stand alone but rather must be a part of a performance management system. A performance management system grounded in the organization's mission-

based strategic intent is necessary. Rushing to measurement and assessment without a clear plan could lead organizational leaders to the wrong metrics and generate another exercise in chasing donor expectations rather than embracing an opportunity to evolve the performance culture of the organization.

Second, the participants sharing the viewpoints expressed how funders expect performance data as a condition of the funding, yet few grantors include funding for administrative and evaluation purposes. The participants who shared viewpoints conveyed a desire for a sense of partnership with their funders as both work in concert toward the same goal of better service delivery for beneficiaries. Characteristics of an effective performance measurement model are partnership and shared goals. To prove effectiveness, leaders of nonprofit organizations need the ability to allocate resources that will enable effectiveness. “Nonprofits are encouraged to collaborate instead of compete, hold onto under-performing staff, accept martyr-like salaries, smile and nod when funders push them in tangential directions and keep quiet” while donors are giving less and expecting more (Edgington, 2010, para. 3). Sector expectations are preventing leaders of nonprofit organizations from creating the impact the sector seeks.

Last, the three perspectives in this study necessitated organization-specific measures linking concretely to the mission rather than ill-fitting and vague standardized metrics to evaluate nonprofit organizational performance. The factors favored outcome measurement as the most effective practice for evaluating a nonprofit’s performance. Outcome measurement is one method for nonprofit organizations to demonstrate impact. When nonprofit leaders have a better understanding of their outcomes and performance data, then there is opportunity for greater collaboration, learning, and increasing the

sector's collective impact. For performance measurement to be effective, it cannot occur in a vacuum.

We need to start thinking about telling the whole story of how the charitable sector functions—how it's funded, how it uses money, how it delivers services, and why this can't be done better by the public sector or the for-profit sector. (Berman, 2013, para. 9)

Nonprofits cannot thrive and prove impact within the restraints of the old paradigm. Stakeholders cannot reduce the human elements of philanthropy to cost-benefit analyses rather than view them as an important civic virtue (Gibson & Dietel, 2010). The new paradigm will require removing overhead ratios and financial efficiency from the lexicon, or at a bare minimum significantly downgrading their importance. It will also require foundations to partner with, mentor, and develop nonprofit organizational capability to undertake such an endeavor.

Solely focusing on the power imbalance in the funder-nonprofit relationship would be an incomplete definition of the problem. Many nonprofit leaders fail to ask the hard questions about their strategy, mission, and performance.

Nonprofits aren't given the chance to have thoughtful and open conversations about the findings in ways that could help them strengthen their own activities. And philanthropists don't have the benefit of getting honest, first-hand perspectives from a broad array of organizations with expertise. (Gibson, 2013, para. 9)

Despite demand for increased transparency, few leaders of nonprofit organizations have embraced the technologies and practices that help to invite and integrate public feedback

and critique into organizational learning, which is likely a function of avoiding overhead and resisting spending on organizational infrastructure and technology.

The current philanthropy paradigm is failing the sector because of unrealistic funder expectations, pressure on nonprofit organizations to conform to these expectations, and misleading reporting and underinvestment in administrative operating expenses (Tierney & Steele, 2011). The desire for change is evident in this study, in addition to the reaction to the overhead myth (Taylor et al., 2013). The three viewpoints discovered in the study demonstrated an understanding that nonprofit leaders need to make changes and be ready to implement the much-needed changes. Other studies have mimicked nonprofit leaders' desire to define effectiveness as outcome accountability. "This sounds like common sense, but if organizational effectiveness is outcome accountability, any meaningful ratings system used to evaluate nonprofits must somehow measure the extent to which organizations achieve their goals" (Mitchell & Schmitz, 2010, para. 6). Rather than viewing financial performance in a vacuum, a growing group of nonprofit leaders is proposing the consideration of other factors, including transparency, governance, leadership, and results (Taylor et al., 2013), which will require that the softer side of philanthropy, the artistic, creative, humanistic side, have as much voice as the scientific side (Gibson & Dietel, 2010).

True change necessitates a shift in the current philanthropy paradigm. Staff and volunteers may have to change behaviors, embrace new ideas, and potentially say goodbye to programs and services that no longer align with the organization's theory of change. Nonetheless, this discomfort and evolution is important because it will enable the organizational leaders to answer questions about the social changes that would

emanate from their mission accomplishment and which variables they should track and measure (Axelrod, 2013).

Change will take patience and a belief that collectively the sector can solve the challenges ahead. “Purpose and perseverance are so important to ultimate success. They must be inflamed, nurtured, and renewed at every turn. And they must be accepted as the bulwark of change” (Light, 2010, *Change by All Means Possible*, para. 6). The collective impact of the sector depends upon this change. Nonprofits strengthen society, increase the common good, and enhance social capital. Changing nonprofit performance expectations is a challenge that stakeholders cannot avoid because of the importance of the sector’s work. The leaders who accept the challenge, those who desire to understand the impact of their work, will be “the ones most likely to make a real difference in the lives of those they serve” (Morino, 2011, p. 22).



## References

- Akhtar-Danesh, N., Baumann, A., & Cordingley, L. (2008). Q methodology in nursing research. *Western Journal of Nursing Research, 30*, 759-773.  
doi:10.1177/0193945907312979
- Akhtar-Danesh, N., Baxter, P., Valaitis, R. K., Stanyon, W., & Sproul, S. (2009). Nurse faculty perceptions of simulation use in nursing education. *Western Journal of Nursing Research, 31*, 312-329. doi:10.1177/0193945908328264
- Alaimo, S. (2008). *Program evaluation capacity for nonprofit human services organizations: An analysis of determining factors* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3342191)
- Alford, J., & Hughes, O. (2008). Public value pragmatism as the next phase of public management. *The American Review of Public Administration, 38*, 130-148.  
doi:10.1177/0275074008314203
- American Institute of Philanthropy. (2010). *Criteria*. Retrieved from <http://www.charitywatch.org/criteria.html>
- Argyris, C. (1976). Single-loop and double-loop models in research on decision making. *Administrative Science Quarterly, 21*, 363-375. doi:10.2307/2391848
- Axelrod, T. (2013, June). Mission envy vs. mission impact. Retrieved from [http://www.guidestar.org/rxa/news/articles/2013/mission-envy-vs-mission-impact.aspx?hq\\_e=el&hq\\_m=2118010&hq\\_l=22&hq\\_v=8334e8a6d0](http://www.guidestar.org/rxa/news/articles/2013/mission-envy-vs-mission-impact.aspx?hq_e=el&hq_m=2118010&hq_l=22&hq_v=8334e8a6d0)
- Babbie, E. R. (2008). *The basics of social research* (4th ed.). Belmont, CA: Thomson Wadsworth.

- Bacon, S., Kimble, R., & Taylor, G. (1989). The wilderness challenge model. In S. Gabel, R. D. Lyman, & S. Prentice-Dunn (Eds.), *Residential and inpatient treatment of children and adolescents* (pp. 115-144). New York, NY: Springer.
- Bailey, M. (2005). Think “results,” not “evaluation.” *The Public Manager*, 34, 8-10.  
Retrieved from <http://www.thepublicmanager.org/>
- Bailin, M. A. (2003). Requestioning, reimagining, and retooling philanthropy. *Nonprofit and Voluntary Sector Quarterly*, 32(4), 635-642. doi:10.1177/0899764003257464
- 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
- Bell-Rose, S. (2004). Using performance metrics to assess impact. In S. Oster (Ed.), *Generating and sustaining nonprofit earned income* (pp. 269-280). San Francisco, CA: Jossey Bass.
- Benjamin, L. M. (2008). Bearing more risk for results: Performance accountability and nonprofit relational work. *Administration & Society*, 39, 959-983.  
doi:10.1177/0095399707309357
- Benjamin, L. M., & Misra, K. (2006). Doing good work: Implications of performance accountability for practice in the nonprofit sector. *International Journal of Rural Management*, 2, 147-162. doi:10.1177/097300520600200202
- Berger, K. (2009). *The worst (and best) way to pick a charity this year*. Retrieved from <http://www.kenscommentary.org/2009/12/worst-and-best-way-to-pick-charity-this.html>

- Berger, K. (2013, April 5). Debating the realities of ranking charities. *Nonprofit Quarterly*. Retrieved from <http://www.nonprofitquarterly.org>
- Berger, K., Penna, R., & Goldberg, S. H. (2010, May). The battle for the soul of the nonprofit sector. *Philadelphia Social Innovations Journal*, (3). Retrieved from <http://www.philasocialinnovations.org/>
- Berman, M. A. (2013, February 28). Sequestering the sector [Web log post]. *Stanford Social Innovation Review*. Retrieved from [http://www.ssireview.org/blog/entry/sequestering\\_the\\_sector?utm\\_source=Enews13\\_02\\_28&utm\\_medium=email&utm\\_content=3&utm\\_campaign=berman](http://www.ssireview.org/blog/entry/sequestering_the_sector?utm_source=Enews13_02_28&utm_medium=email&utm_content=3&utm_campaign=berman)
- Better Business Bureau Wise Giving Alliance. (2010). *Standards for charity accountability*. Retrieved from <http://www.bbb.org/us/Charity-Standards/>
- Bibb-Binder, E., & Kramer, K. (2013, June). Facing the future: How successful nonprofits link strategy to leadership development. Retrieved from <http://www.bridgespan.org/Publications-and-Tools/Career-Professional-Development/Develop-My-Staff/Facing-the-Future.aspx#.UgVRfFPB2c9>
- Birdsong, M., & Ni, P. (2012, June 24). To build better programs, charities must listen to the people they help. *Chronicle of Philanthropy*. Retrieved from <http://philanthropy.com/article/CharitiesGrant-Makers/132485/>
- Bladt, J., & Filbin, B. (2013, March 4). Know the difference between your data and your metrics [Web log post]. Retrieved from [http://blogs.hbr.org/cs/2013/03/know\\_the\\_difference\\_between\\_yo.html](http://blogs.hbr.org/cs/2013/03/know_the_difference_between_yo.html)

- Breckenridge, J., & Jones, D. (2009). Demystifying theoretical sampling in grounded theory research. *The Grounded Theory Review*, 8(8), 114-126. Retrieved from <http://www.groundedtheoryreview.com/>
- Brock, A., Buteau, E., & Herring, A. (2012). *Room for improvement: Foundations' support of nonprofit performance assessment*. Cambridge, MA: Center for Effective Philanthropy. Retrieved from [http://www.jfunders.org/sites/default/files/docs/resources/room\\_for\\_improvement.pdf](http://www.jfunders.org/sites/default/files/docs/resources/room_for_improvement.pdf)
- Brown, L. D. (2008). *Creating credibility: Legitimacy and accountability for transnational civil society*. Sterling, VA: Kumarian Press.
- Brown, L.D. & Kalegaonkar, A. (2002). Support organizations and the evolution of the NGO sector. *Nonprofit and Voluntary Sector Quarterly*, 31(2), 231-258.  
doi:10.1177/0899764002312004
- Brown, S. R. (1980). *Political subjectivity: Applications of Q methodology in political science*. New Haven, CT: Yale University Press.
- Buchman, T. (2009). *After the funding: Perceptions and realities of health-care emergency preparedness* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3381822)
- Buckmaster, N. (1999). Associations between outcome measurement, accountability and learning for non-profit organisations. *The International Journal of Public Sector Management*, 12, 186-197. doi:10.1108/09513559910263499
- Burr, V. (1995). *An introduction to social constructionism*. London, England: Routledge.  
doi:10.4324/9780203299968

- Callen, J., Klein, A., & Tinkleman, D. (2003). Board composition, committees, and organizational efficiency: The case of nonprofits. *Nonprofit and Voluntary Sector Quarterly*, 32, 493-521. doi:10.1177/0899764003257462
- Cameron, K. S., & Whetten, D. A. (1983). Some conclusions about organizational effectiveness. In K. S. Cameron & D. A. Whetten (Eds.), *Organizational effectiveness: A comparison of multiple methods* (pp. 261-277). New York, NY: Academic Press.
- Carman, J. G. (2007). Evaluation practice among community-based organizations: Research into the reality. *American Journal of Evaluation*, 28, 60-75. doi:10.1177/1098214006296245
- Carman, J. G. (2009). Nonprofits, funders, and evaluation: Accountability in action. *The American Review of Public Administration*, 39, 374-390. doi:10.1177/0275074008320190
- 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., & Fredericks, K. A. (2010). Evaluation capacity and nonprofit organizations: Is the glass half-empty or half-full? *American Journal of Evaluation*, 31, 84-104. doi:10.1177/1098214009352361
- Carson, E. D. (2000). On foundations and outcome evaluation. *Nonprofit and Voluntary Sector Quarterly*, 29, 479-481. doi:10.1177/0899764000293008

- Carttar, P. (2013, March 8). Why don't the best nonprofits grow [Web log post]?  
Retrieved from  
[http://blogs.hbr.org/cs/2013/03/social\\_enterprises\\_cant\\_grow\\_w.html](http://blogs.hbr.org/cs/2013/03/social_enterprises_cant_grow_w.html)
- Cohen, L., Manion, L., Morrison, K., & Morrison, K. R. B. (2007). *Research methods in education*. New York, NY: Routledge.
- Choi, Y. J. (2010). The ideal female body image: A Q approach to the third-person effect. *Operant Subjectivity, 33*, 129-148.
- Colby, S., Stone, N., & Carttar, P. (2004). Zeroing in on impact. *Stanford Social Innovation Review, 2*(2). Retrieved from <http://www.ssireview.org/>
- Congressional Budget Office. (2011). *Options for changing the tax treatment of charitable giving* (Pub. No. 4030). Washington, DC: The Congress of the United States.
- Connolly, T., Conlon, E. J., & Deutsch, S. J. (1980). Organizational effectiveness: A multiple-constituency approach. *Academy of Management Review, 5*, 211-217.  
Retrieved from <http://www.aom.pace.edu/amr/>
- Cooper, D. R., & Schindler, P. S. (2008). *Business research methods* (10th ed.). Boston, MA: McGraw-Hill Irwin.
- Cooperrider, D. L., & Whitney, D. K. (2005). *Appreciative inquiry: A positive revolution in change*. San Francisco, CA: Berrett-Koehler.
- Crawford, K. (2013, April 1). The hidden biases in big data [Web log post]. Retrieved from [http://blogs.hbr.org/cs/2013/04/the\\_hidden\\_biases\\_in\\_big\\_data.html](http://blogs.hbr.org/cs/2013/04/the_hidden_biases_in_big_data.html)

- Creswell, J. W. (2005). *Educational research: Planning, conducting and evaluating quantitative and qualitative research* (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage.
- Crutchfield, L. R., & Grant, H. M. (2008). *Forces for good: The six practices of high-impact nonprofits*. San Francisco, CA: Jossey-Bass.
- Cunliffe, A. L. (2004). On becoming a critically reflexive practitioner. *Journal of Management Education*, 28, 407-426. doi:10.1177/1052562904264440
- Davenport, J., & Gardiner, P. D. (2007). Performance management in the not-for-profit sector with reference to the National Trust for Scotland. *Total Quality Management*, 18, 303-311. doi:10.1080/14783360601152517
- de Graaf, G., & van Exel, J. (2008). Using Q methodology in administrative ethics. *Public Integrity*, 11, 63-78. doi:10.2753/PIN1099-9922110104
- Denhardt, R. B., Denhardt, J. V., & Aristigueta, M. P. (2009). *Managing human behavior in public and nonprofit organizations*. Thousand Oaks, CA: Sage.
- De Vita, C. J., & Fleming, C. (Eds.). (2001). *Building capacity in nonprofit organizations*. Washington, DC: The Urban Institute.
- Donaldson, S. I. (2002). Overcoming our negative reputation: Evaluation becomes known as a helping profession. *American Journal of Evaluation*, 22(3), 355-361.
- Retrieved from <http://aje.sagepub.com/>

- Dziopa, F., & Ahern, K. (2011). A systematic literature review of the applications of Q-technique and its methodology. *Methodology*, 7(2), 39-55. doi:10.1027/1614-2241/a000021
- Easterling, D. (2000). Using outcome evaluation to guide grantmaking: Theory, reality, and possibilities. *Nonprofit and Voluntary Sector Quarterly*, 29(3), 482-486. doi:10.1177/0899764000293009
- Easterly, L., & Miesing, P. (2009). NGOs, social venturing, and community citizenship behavior. *Business Society*, 48, 538-565. doi:10.1177/0007650308317385
- Ebrahim, A. (2005). Accountability myopia: Losing sight of organizational learning. *Nonprofit and Voluntary Sector Quarterly*, 34, 56-87. doi:10.1177/0899764004269430
- Ebrahim, A. (2009). Placing the normative logics of accountability in “thick” perspective. *American Behavioral Scientist*, 52, 885-904. doi:10.1177/0002764208327664
- Ebrahim, A., & Weisband, E. (Eds.). (2007). *Global accountabilities: Participation, pluralism, and public ethics*. Cambridge, England: Cambridge University Press.
- Edgington, N. (2010, March 24). The perils of nice. Retrieved from [http://socialentrepreneurship.change.org/blog/view/the\\_perils\\_of\\_nice](http://socialentrepreneurship.change.org/blog/view/the_perils_of_nice)
- Eisinger, P. (2002). Organizational capacity and organizational effectiveness among street-level food assistance programs. *Nonprofit and Voluntary Sector Quarterly*, 31, 115-130. doi:10.1177/0899764002311005
- Emerson, J., Wachowiz, J., & Chun, S. (2000). *Social return on investment: Exploring aspects of value creation in the nonprofit sector*. San Francisco, CA: The Roberts Foundation.



- Fine, A. H., Thayer, C. E., & Coghlan, A. (2000). Program evaluation practice in the nonprofit sector. *Nonprofit Management and Leadership, 10*, 331-339.  
doi:10.1002/nml.10309
- 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
- Ford, M. W., & Evans, J. R. (2002). Baldrige assessment and organizational learning: The need for change management. *Quality Management Journal, 47*, 153-156.  
Retrieved from <http://asq.org/pub/qmj/>
- Foster, M. J., & Lock, A. R. (1990). Factoring effectiveness factors! *Journal of the Operational Research Society, 41*, 111-117. Retrieved from <http://www.palgrave-journals.com/jors/index.html>
- Foster, W. L., Kim, P., & Christiansen, B. (2009). Ten nonprofit funding models. *Stanford Social Innovation Review, 7*(2). Retrieved from <http://www.ssireview.org/>
- Fowler, F. J. (2009). *Survey research methods* (4th ed.). Thousand Oaks, CA: Sage.
- Fram, E. H., & Talley, J. L. (2012, July 24). Using imperfect metrics well: Tracking progress and driving change. *Management: Nonprofit Quarterly*. Retrieved from <http://nonprofitquarterly.org>
- Fredericksen, P., & London, R. (2000). Disconnect in the hollow state: The pivotal role of organizational capacity in community-based development organizations. *Public Administration Review, 60*, 230-239. doi:10.1111/0033-3352.00083

- Froelich, K. A. (1999). Diversification of revenue strategies: Evolving resource dependence in nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly*, 28, 246-268. doi:10.1177/0899764099283002
- Gelo, O., Braakmann, D., & Benetka, G. (2008). Quantitative and qualitative research: Beyond the debate. *Integrative Psychological & Behavioral Science*, 42, 266-290. doi:10.1007/s12124-008-9078-3
- Giannoulis, C., Botetzagias, I., & Skanavis, C. (2010). Newspaper reporters' priorities and beliefs about environmental journalism: An application of Q methodology. *Science Communication*, 32, 425-466. doi:10.1177/1075547010364927
- Gibson, C. (2013, June 16). Open debate, not politeness, is what drives nonprofit innovation. *Community: The Chronicle of Philanthropy*. Retrieved from <http://philanthropy.com>
- Gibson, C., & Dietel, W. M. (2010). What do donors want? *Nonprofit Quarterly*, 17(3). Retrieved from <http://nonprofitquarterly.org/>
- Goggins-Gregory, A., & Howard, D. (2009). The nonprofit starvation cycle. *Stanford Social Innovation Review*, 7(4). Retrieved from <http://www.ssireview.org/>
- Govindarajan, V. (2012, October 22). The timeless strategic value of unrealistic goals [Web log post]. Retrieved from <http://blogs.hbr.org/govindarajan/2012/10/the-timeless-strategic-value-of-unrealistic-goals.html>
- Grant Thornton. (2008). *New Form 990 focuses on your governance, and more*. Retrieved from <http://www.grantthornton.com/staticfiles/GTCom/files/Industries/NotForProfit/NFP%20Bulletin%20-%20issue%201%20-%203.pdf>

- Greenwood, R., Oliver, C., Suddaby, R., & Sahlin-Andersson, K. (2008). *The Sage handbook of organizational institutionalism*. Thousand Oaks, CA: Sage.
- Groves, R. (2013, April 4). What moneyball didn't teach you about measurement [Web log post]. Retrieved from [http://philanthropy.com/blogs/measuring-up/what-moneyball-didnt-teach-you-about-measurement/87?cid=pt&utm\\_source=pt&utm\\_medium=en](http://philanthropy.com/blogs/measuring-up/what-moneyball-didnt-teach-you-about-measurement/87?cid=pt&utm_source=pt&utm_medium=en)
- Hamel, G., & Prahalad, C.K. (1989, May-June). Strategic intent. *Harvard Business Review*, 63-76.
- Hathcoat, J. D., & Montgomery, D. (2010). Second-order structure of academic and religious personal epistemologies. *Operant Subjectivity*, 33, 26-50. Retrieved from <http://operantsubjectivity.org/os/about>
- Havens, J. J., & Schervish, P. G. (2010, January/February). Center on Wealth and Philanthropy individual giving model: Forecast for 2009. *Advancing Philanthropy*, 30-35. Retrieved from <http://www.afpnet.org/publications/issuelist.cfm>
- Heimovics, R. D., Herman, R. D., & Coughlin, C. L. J. (1993). Executive leadership and resource dependence in nonprofit organizations: A frame analysis. *Public Administration Review*, 53, 419-427. doi:10.2307/976342
- Henderson, D. A., Chase, B. W., & Woodson, B. M. (2002). Performance measures for NPOs. *Journal of Accountancy*, 193, 63-67. Retrieved from <http://www.journalofaccountancy.com/>

- Herman, R. D. (1992). Nonprofit organization effectiveness: At what, for whom, according to whom? *Nonprofit and Voluntary Sector Quarterly*, 21(4), 411-416.  
doi:10.1177/089976409202100407
- Herman, R. D., & Renz, D. O. (1997). Multiple constituencies and the social construction of nonprofit organization effectiveness. *Nonprofit and Voluntary Sector Quarterly*, 26, 185-206. doi:10.1177/0899764097262006
- Herman, R. D., & Renz, D. O. (1999). Theses on nonprofit organizational effectiveness. *Nonprofit and Voluntary Sector Quarterly*, 28, 107-126.  
doi:10.1177/0899764099282001
- Herman, R. D., & Renz, D. O. (2004). Doing things right: Effectiveness in local nonprofit organizations, a panel study. *Public Administration Review*, 64, 694-704.  
doi:10.1111/j.1540-6210.2004.00416.x
- Herman, R. D., & Renz, D. O. (2008). Advancing nonprofit organizational effectiveness research and theory: Nine theses. *Nonprofit Management & Leadership*, 18, 399-416. doi:10.1002/nml.195
- Hetrick, M. J. (2004). *Performance of nonprofit human service organizations affiliated with the United Way*. (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3135998)
- Hoefer, R. (2000). Accountability in action? Program evaluation in nonprofit human service agencies. *Nonprofit Management and Leadership*, 11, 167-177.  
doi:10.1002/nml.11203
- Hoffman, S. (2006, November 21). For U.S. charities, a crisis of trust. *NBC News*. Retrieved from <http://www.msnbc.msn.com/id/15753760/>

- Holt, G. E. (2006). SOX: “Best practices” or too much accountability. *The Bottom Line*, 19(3), 139-145. doi:10.1108/08880450610682554
- Holton, E. F. (1999). Performance domains and their boundaries. *Advances in Developing Human Resources*, 1(1), 26-37. doi: 10.1177/152342239900100103
- Hope Consulting. (2011). *Money for good II: Full report*. Retrieved from <http://www.hopeconsulting.us/moneyforgood>
- Huang, J., Buchanan, P., & Buteau, E. (2006). *In search of impact*. Cambridge, MA: Center for Effective Philanthropy.
- Hundley, K., & Taggart, K. (2013, June 6). America’s 50 worst charities rake in nearly \$1 billion for corporate fundraisers. *Tampa Bay Times*. Retrieved from <http://www.tampabay.com/>
- Independent Sector. (2011). *Word on Washington*. Retrieved from [http://independentsector.org/word\\_on\\_washington](http://independentsector.org/word_on_washington)
- Independent Sector. (2012a). *Scope of the nonprofit sector*. Retrieved from [http://independentsector.org/scope\\_of\\_the\\_sector](http://independentsector.org/scope_of_the_sector)
- Independent Sector. (2012b). *The sector’s economic impact*. Retrieved from [http://independentsector.org/economic\\_role](http://independentsector.org/economic_role)
- Ingram, S. H. (2009). *Nonprofit governance: The view from the IRS*. Retrieved from [http://www.irs.gov/pub/irs-tege/ingram\\_\\_gtown\\_\\_governance\\_062309.pdf](http://www.irs.gov/pub/irs-tege/ingram__gtown__governance_062309.pdf)
- Jagpal, N., & Laskowski, K. (2013). Three ways to boost core support [Web log post]. *Foundations: Stanford Social Innovation Review*. Retrieved from [http://www.ssireview.org/blog/entry/three\\_ways\\_to\\_boost\\_core\\_support?utm\\_source=Enews&utm\\_medium=email&utm\\_content=3&utm\\_campaign=From\\_Blog](http://www.ssireview.org/blog/entry/three_ways_to_boost_core_support?utm_source=Enews&utm_medium=email&utm_content=3&utm_campaign=From_Blog)

- Jedeloo, S., van Staa, A., Latour, J. M., & van Exel, N. J. A. (2010). Preferences for health care and self-management among Dutch adolescents with chronic conditions: A Q methodological investigation. *International Journal of Nursing Studies, 47*, 593-603. doi:10.1016/j.ijnurstu.2009.10.006
- Judge, W. Q. (1994). Correlates of organizational effectiveness: A multilevel analysis of a multidimensional outcome. *Journal of Business Ethics, 13*, 1-10.  
doi:10.1007/BF00877149
- 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. (1996). Using the balanced scorecard as a strategic management system. *Harvard Business Review, 74*, 75-85. Retrieved from <http://hbr.org/>
- Katz, D. & Kahn. R.L. (1978). *The social psychology of organizations* (2nd ed.). Hoboken, NJ: John Wiley & Sons.
- Kelly, D., & Lewis, A. (2009). Human service sector nonprofit organization's social impact. *Business Strategy Series, 10*(6), 374-382.  
doi:10.1108/17515630911005664
- Khosrowpour, M. (2002). *Advanced topics in information resources management*. Hershey, PA: Idea Group.
- Kluger, M.P. (n.d.). The program evaluation grid: A planning and assessment tool for nonprofit organizations. *Administration in Social Work, 30*(1), 33-44. Retrieved from <http://www.tandf.co.uk/journals/titles/03643107.asp>

- Kohn, A. (1986). *No contest: The case against competition*. Boston, MA: Houghton Mifflin.
- Kondalkar, V. G. (2009). *Organization effectiveness and change management*. New Delhi, India: PHI Learning.
- Kramer, K., & Stid, D. (n.d.). The effective organization: Five questions to translate leadership into strong management. Retrieved from <http://www.bridgespan.org/Publications-and-Tools/Organizational-Effectiveness/The-Effective-Organization-Five-Questions-to-Trans.aspx#UgU4rIPB2c9>
- Kuhn, T. (1996). *The structure of scientific revolutions* (3rd ed.). Chicago, IL: University of Chicago Press.
- Lampkin, L. M., & Boris, E. T. (2002). Nonprofit organization data: What we have and what we need. *American Behavioral Scientist*, *45*, 1675-1716.  
doi:10.1177/0007650308317385
- Larkin, R. (2013, July 2). Using outcomes to measure nonprofit success. *Management: Nonprofit Quarterly*. Retrieved from <http://www.nonprofitquarterly.org/management/22549-using-outcomes-to-measure-nonprofit-success.html>
- Leat, D. (2006). Grantmaking foundations and performance measurement. *Public Policy and Administration*, *21*(3), 25-37. doi:10.1177/095207670602100303
- Lebas, M., & Euske, K. (2002). A conceptual and operational delineation of performance. In A. Neely (Ed.), *Business performance measurement* (pp. 63-78). Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511753695.006
- Lecy, J., Schmitz, H. P., & Swedlund, H. (n.d.). *Mapping research traditions: Assessing the effectiveness of non-governmental and non-for-profit organizations* (Working

- Paper). Retrieved from [http://www1.maxwell.syr.edu/uploadedFiles/Moynihan/tngo/NGOEffectivenessLitReview\\_Haley\\_12-15-09.pdf](http://www1.maxwell.syr.edu/uploadedFiles/Moynihan/tngo/NGOEffectivenessLitReview_Haley_12-15-09.pdf)
- LeRoux, K. (2009). Managing stakeholder demands: Balancing responsiveness to clients and funding agents in nonprofit social service organizations. *Administration Society, 41*, 158-184. doi:10.1177/0095399709332298
- Letts, C., Ryan, W. P., & Grossman, A. (1999). *High performance nonprofit organizations: Managing upstream for greater impact*. New York, NY: Wiley.
- Lewin, A. Y., & Minton, J. W. (1986). Determining organizational effectiveness: Another look, and an agenda for research. *Management Science, 32*, 514-538. doi:10.1287/mnsc.32.5.514
- Lewis, D. (2001). *The management of non-governmental development organizations: An introduction*. Oxford, England: Routledge. doi:10.4324/9780203002162
- Light, P. (2010, Winter). Driving social change. *Nonprofit Quarterly, 17*(4). Retrieved from <http://www.nonprofitquarterly.org/>
- Light, P. C. (2002). *Pathways to nonprofit excellence*. Washington, DC: Brookings Institution Press.
- Lindenberg, M. (2001). Are we at the cutting edge or the blunt edge? *Nonprofit Management & Leadership, 11*, 247-270. doi:10.1002/nml.11302
- Lovett, K. (2012, May 16). Gov. Andrew Cuomo pushes plan to cap pay of nonprofit execs at \$199,000. *New York Daily News*. Retrieved from <http://www.nydailynews.com/>



- Lowell, S., Silverman, L., & Taliento, L. (2001). Not-for-profit management: The gift that keeps on giving. *The McKinsey Quarterly*, 1, 147-155. Retrieved from <http://www.mckinseyquarterly.com/home.aspx>
- Lowell, S., Trelstad, B., & Meehan, B. (2005). The ratings game. *Stanford Social Innovation Review*, 3(2), 39-45. Retrieved from <http://www.ssireview.org/>
- Lumley, T. (2013, July 10). Raising the bar on nonprofit impact measurement. *Nonprofits (SSRI)*. Retrieved from [http://www.ssireview.org/blog/entry/raising\\_the\\_bar\\_on\\_nonprofits\\_impact\\_measurement?utm\\_source=Enews&utm\\_medium=email&utm\\_content=3&utm\\_campaign=From\\_Blog](http://www.ssireview.org/blog/entry/raising_the_bar_on_nonprofits_impact_measurement?utm_source=Enews&utm_medium=email&utm_content=3&utm_campaign=From_Blog)
- Mayne, J. (2007). Challenges and lessons in implementing results-based management. *Evaluation*, 13, 87-109. doi:10.1177/1356389007073683
- McCambridge, R. (2012a, August 15). External influences on nonprofit management: A wide-angle view. *Management: Nonprofit Quarterly*. Retrieved from <http://nonprofitquarterly.org/>
- McCambridge, R. (2012b, May 21). New study finds non-reporting of fundraising expense widespread: But some states much worse than others. *Management: Nonprofit Quarterly*. Retrieved from <http://nonprofitquarterly.org/>
- McKeown, B., & Thomas, D. (2013). *Q methodology: Quantitative applications in the social sciences* (2nd ed.) Newbury Park, CA: Sage.
- McKeown, M., Hinks, M., Stowell-Smith, M., Mercer, D., & Forster, J. (1999). Q methodology, risk training and quality management. *International Journal of*

*Health Care Quality Assurance*, 12(6), 254-266.

doi:10.1108/09526869910291823

Mensah, Y., Lam, K., & Werner, R. (2008). An approach to evaluating the relative effectiveness in non-profit institutions. *Journal of Public Budgeting, Accounting & Financial Management*, 20, 324-354. Retrieved from <http://pracademics.com/jpbafm.html>

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(1), 83-94. doi: 10.1177/1098214005283185

Miller-McLemore, M. (2012, July). Nonprofit assessment: Richer perspectives. *GuideStar*. <http://www.guidestar.org/>

Mittendorf, B. (2013, June 5). Want charities to be evaluated based on impact? Be careful what you wish for. *Management: Nonprofit Quarterly*. Retrieved from <http://nonprofitquarterly.org/>

Molnár, M. (2008). The accountability paradigm: Standards of excellence. *Public Management Review*, 10, 127-137. doi:10.1080/14719030701763245

Montgomery, D. (2010). Q applied to educational issues. *Operant Subjectivity*, 33, 1-2. Retrieved from <http://operantsubjectivity.org/os/about>

Moore, M. H. (2000). Managing for value: Organizational strategy in for-profit, nonprofit, and governmental organizations. *Nonprofit and Voluntary Sector Quarterly*, 29(1), 183-208. doi:10.1177/089976400773746391

Moore, M. H. (2003). *The public value scorecard: A rejoinder and an alternative to "Strategic performance measurement and management in non-profit*

- organizations*” by Robert Kaplan (Working Paper No. 18). Cambridge, MA: Hauser Center for Nonprofit Organizations. doi:10.2139/ssrn.402880
- Morariu, J., Athanasiades, K., & Emery, A. (2012). State of evaluation 2012: Evaluation practice and capacity in the nonprofit sector. Retrieved from [http://www.innonet.org/client\\_docs/innonet-state-of-evaluation-2012.pdf](http://www.innonet.org/client_docs/innonet-state-of-evaluation-2012.pdf)
- Morgan, W. M. (2006). *Measuring performance in nonprofit organizations: A stochastic frontier approach* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3217979)
- Morino, M. (2011). *Leap of reason: Managing to outcomes in an era of scarcity*. Washington, DC: Venture Philanthropy Partners.
- Morley, E., Vinson, E., & Hatry, H. P. (2001). *Outcome measurement in nonprofit organizations: Current practices and recommendations* (pp. 1-10). Washington, DC: Independent Sector & The Urban Institute.
- Morrow, P. C. (1982). Explorations in macro communication behaviour: The effects of organizational feedback on organizational effectiveness. *Journal of Management Studies*, 19, 437-446. Retrieved from <http://onlinelibrary.wiley.com/journal/10.1111/%28ISSN%291467-6486>
- Moxham, C. (2009). Performance measurement: Examining the applicability of the existing body of knowledge to nonprofit organisations. *International Journal of Operations & Production Management*, 29(7), 740-763. doi:10.1108/01443570910971405
- National Institute of Standards and Technology. (2010). *Self-assessing your organization*. Retrieved from <http://www.nist.gov/baldrige/enter/self.cfm>

- Neuman, W. L. (2006). *Social research methods: Qualitative and quantitative approaches*. Boston, MA: Pearson/Allyn and Bacon.
- O'Donovan, D., & Rimland-Flower, N. (2013, January 10). The strategic plan is dead: Long live strategy [Web log post]. *Nonprofit Management: Stanford Social Innovation Review*. Retrieved from [http://www.ssireview.org/blog/entry/the\\_strategic\\_plan\\_is\\_dead.\\_long\\_live\\_strategy?utm\\_source=Enews&utm\\_medium=email&utm\\_campaign=ten\\_gifts](http://www.ssireview.org/blog/entry/the_strategic_plan_is_dead._long_live_strategy?utm_source=Enews&utm_medium=email&utm_campaign=ten_gifts)
- O'Flanagan, M., Harold, J., & Brest, P. (2008). *The nonprofit marketplace: Bridging the information gap in philanthropy*. Menlo Park, CA: The William and Flora Hewlett Foundation and McKinsey & Company.
- Ospina, S., Diaz, W., & O'Sullivan, J. F. (2002). Negotiating accountability: Managerial lessons from identity-based nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly*, 31, 5-31. doi:10.1177/0899764002311001
- Osterkiil, C. (2010). Second-order structure of academic and religious personal epistemologies. *Operant Subjectivity*, 33, 51-67. Retrieved from <http://operantsubjectivity.org/os/about>
- Ott, J. S., & Shafritz, J. M. (1994). Toward a definition of organizational incompetence: A neglected variable in organization theory. *Public Administration Review*, 54, 370-377. Retrieved from <http://www.blackwellpublishing.com/journal.asp?ref=0033-3352>
- Pallotta, D. (2008). *Uncharitable: How restraints on nonprofits undermine their potential*. Medford, MA: Tufts University Press.

- Peck, R. (2010). Introduction to the Q-block symposium. *Operant Subjectivity*, 34, 1-5.  
Retrieved from <http://operantsubjectivity.org/os/about>
- Penrod, J., Preston, D. B., Cain, R. E., & Starks, M. T. (2003). A discussion of chain referral as a method of sampling hard-to-reach populations. *Journal of Transcultural Nursing*, 14, 100-107. doi:10.1177/1043659602250614
- Pfeffer, J. S., & Salancik, G. (1978). *The external control of organizations: A resource dependence perspective*. New York, NY: Harper & Row.
- Polonsky, M. J., & Grau, S. L. (2008). Evaluating the social value of charitable organizations: A conceptual foundation. *Journal of Macromarketing*, 28, 130-140. doi:10.1177/0276146708314585
- 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 (PAQS). *American Journal of Evaluation*, 21(1), 15-26. doi:10.1177/109821400002100102
- Porter, M. E., & Kramer, M. R. (1999). Philanthropy's new agenda: Creating value. *Harvard Business Review*, 77(6) 121-130. Retrieved from <http://hbr.org/>
- Price, J. L. (1968). The study of organizational effectiveness. *The Sociological Quarterly*, 13, 3-15. doi:10.1111/j.1533-8525.1972.tb00788.x
- Pratt, J. (2013, April 17). Flaws in the social impact bond/pay for success craze. *Policy/Social Context: Nonprofit Quarterly*. Retrieved from <http://nonprofitquarterly.org/>

- Preston, C. (2011). A potential \$15-billion windfall for effective nonprofits. Retrieved from [http://philanthropy.com/blogs/the-giveaway/a-potential-15-billion-windfall-for-effective-nonprofits/1059?sid=pt&utm\\_source=pt&utm\\_medium=en](http://philanthropy.com/blogs/the-giveaway/a-potential-15-billion-windfall-for-effective-nonprofits/1059?sid=pt&utm_source=pt&utm_medium=en)
- Preston, C. (2012, September 13). Nonprofits are dissatisfied with foundations' evaluation efforts. Retrieved from [http://philanthropy.com/blogs/the-giveaway/nonprofits-are-dissatisfied-with-foundations-evaluation-efforts/3194?cid=pt&utm\\_source=pt&utm\\_medium=en](http://philanthropy.com/blogs/the-giveaway/nonprofits-are-dissatisfied-with-foundations-evaluation-efforts/3194?cid=pt&utm_source=pt&utm_medium=en)
- Provan, K. G., & Stewart, D. W. (1982). Measuring organizational effectiveness in the not-for-profit sector. *Academy of Management Proceedings*, 339-343. doi: 10.5465/AMBPP.1982.4976777
- Quinn, R. E., & Rohrbaugh, J. (1983). A spatial model of effectiveness criteria: Towards a competing values approach to organizational analysis. *Management Science*, 29, 363-377. Retrieved from <http://www.informs.org/Pubs/ManSci>
- Ritchie, W. J., Anthony, W. P., & Rubens, A. J. (2004). Individual executive characteristics: Explaining the divergence between perceptual and financial measures in nonprofit organizations. *Journal of Business Ethics*, 53, 267-281. Retrieved from <http://www.springer.com/social+sciences/applied+ethics/journal/10551>
- Ritchie, W. J., & Kolodinsky, R. W. (2003). Nonprofit organization financial performance measurement. *Nonprofit Management & Leadership*, 13, 367-382. Retrieved from <http://onlinelibrary.wiley.com/journal/10.1002/%28ISSN%291542-7854>

- Rosenman, M. (2013a). Nonprofits need to denounce bad fundraising practices. *The Chronicle of Philanthropy*. Retrieved from <http://philanthropy.com/>
- Rosenman, M. (2013b, February 22). Scorecards and what we value. *Huffington Post*. Retrieved from <http://www.huffingtonpost.com/>
- Rubin, A., & Babbie, E. R. (2010). *Research methods for social work*. Belmont, CA: Cengage Learning.
- Salkind, N. J. (2008). *Exploring research* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Sarup, M. (1993). *An introductory guide to post-structuralism and postmodernism*. Harlow, England: Longman.
- Sawhill, J. C., & Williamson, D. (2001). Mission impossible? Measuring success in nonprofit organizations. *Nonprofit Management & Leadership*, 11, 371-387. Retrieved from <http://onlinelibrary.wiley.com/journal/10.1002/%28ISSN%291542-7854>
- Schambra, W. (2013a, April 5). Charity Navigator 3.0: The empirical empire's death star? *Philanthropy: Nonprofit Quarterly*. Retrieved from <http://www.nonprofitquarterly.org/>
- Schambra, W. (2013b, June 10). Why can't we get over overhead? *Management: Nonprofit Quarterly*. Retrieved from <http://www.nonprofitquarterly.org/>
- Schermerhorn, D. J. R., Hunt, D. J. G., & Osborn, D. R. N. (2008). *Organizational behavior* (10th ed.). New York, NY: Wiley.
- Schmolck, P. (2002). PQMethod software. Retrieved from <http://www.lrz.de/~schmolck/qmethod/>

- Schreifels, J. (2013, July 8). About high fundraising costs: It's complicated. *The Chronicle of Philanthropy*. Retrieved from <http://philanthropy.com/>
- Scott, W. R., & Davis, G. F. (2007). *Organizations and organizing: Rational, natural, and open systems perspectives*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Senge, P. (1990). *The fifth discipline: The art and science of the learning organization*. New York, NY: Currency Doubleday.
- Sheehan, R. M., Jr. (1996). Mission accomplishment as philanthropic organization effectiveness: Key findings from the excellence in philanthropy project. *Nonprofit and Voluntary Sector Quarterly*, 25, 110-123. Retrieved from <http://nvs.sagepub.com/>
- Sherry, C. E. (2013). Stop funding duplicative projects. *Stanford Social Innovation Review*, 11(3). Retrieved from <http://www.ssireview.org/>
- Shilbury, D., & Moore, K. A. (2006). A study of organizational effectiveness for national Olympic sporting organizations. *Nonprofit and Voluntary Sector Quarterly*, 35, 5-38. doi:10.1177/0899764005279512
- Sickler, J., Fraser, J., Webler, T., Reiss, D., Boyle, P., Lyn, H., . . . Gruber, S. (2006). Social narratives surrounding dolphins: Q method study. *Society & Animals*, 14, 351-382. doi:10.1163/156853006778882457
- Sloan, M. F. (2009). The effects of nonprofit accountability ratings on donor behavior. *Nonprofit and Voluntary Sector Quarterly*, 38, 220-236. doi:10.1177/0899764008316470
- Smith, G. S. (1988). Performance evaluation for nonprofits. *Nonprofit World*, 6, 24-26. Retrieved from <http://www.snpo.org/publications/nonprofitworld.php>



- Smith, S. R. (2010). Nonprofits and public administration: reconciling performance management and citizen engagement. *The American Review of Public Administration*, 40, 129-152. doi:10.1177/0275074009358452
- Smith-Milway, K., & Goggins-Gregory, A. (2013, February 22). Making organizational learning stick: How to set your knowledge goals and blend technology and people processes to reach them. *Management: Nonprofit Quarterly*. Retrieved from <http://www.nonprofitquarterly.org/>
- 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. Retrieved from <http://nvs.sagepub.com/>
- Spitzer, D. R. (2007). *Transforming performance measurement: Rethinking the way we measure and drive organizational success*. New York, NY: AMACOM Books.
- Spreen, M., & Zwaagstra, R. (1994). Personal network sampling, outdegree analysis and multilevel analysis: Introducing the network concept in studies of hidden populations. *International Sociology*, 9, 475-491.  
doi:10.1177/026858094009004006
- Starkey, K., Tempest, S., & McKinlay, A. (2004). *How organizations learn: Managing the search for knowledge*. London, England: Thomson Learning.
- Starnes, B. J. (2000). Achieving competitive advantage through the application of open systems theory and the development of strategic alliances: A guide for managers of nonprofit organizations. *Journal of Nonprofit & Public Sector Marketing*, 8(2), 15-27. Retrieved from <http://www.tandfonline.com/loi/wnon20>

- Starr, K. (2011, August 3). Just give 'em the money: The power and pleasure of unrestricted funding [Web log post]. *Foundations: Stanford Social Innovation Review*. Retrieved from [http://www.ssireview.org/blog/entry/just\\_give\\_em\\_the\\_money\\_the\\_power\\_and\\_pleasure\\_of\\_unrestricted\\_funding?utm\\_source=Enews&utm\\_medium=email&utm\\_campaign=ten\\_gifts](http://www.ssireview.org/blog/entry/just_give_em_the_money_the_power_and_pleasure_of_unrestricted_funding?utm_source=Enews&utm_medium=email&utm_campaign=ten_gifts)
- Steelman, T. A., & Maguire, L. A. (1999). Understanding participant perspectives: Q methodology in national forest management. *Journal of Policy Analysis and Management*, 18, 361-388. Retrieved from <http://onlinelibrary.wiley.com/journal/10.1002/%28ISSN%291520-6688>
- Steers, R. M. (1977). *Organizational effectiveness: A behavioral view*. Santa Monica, CA: Goodyear.
- Steinem, G. (2012). *Revolution from within: A book of self-esteem* [E-reader version]. New York, NY: Open Road Integrated Media.
- Stephenson, W. (1953). *The study of behavior: Q-technique and its methodology*. Chicago, IL: University of Chicago Press.
- Stevens, S. (2013). Taking on Oregon's rogue charities. *Portland Business Journal*. Retrieved from <http://www.bizjournals.com/portland/>
- Strichman, N., Bickel, W. E., & Marshood, F. (2008). Adaptive capacity in Israeli social change nonprofits. *Nonprofit and Voluntary Sector Quarterly*, 37, 224-248. doi:10.1177/0899764007304462
- Sussman, C. (2003). Making change: How to build adaptive capacity. *The Nonprofit Quarterly*, 10(4), 19-24. Retrieved from <http://www.nonprofitquarterly.org/>

- Swetnam, K. G. (2010). Dropout and academic achievement perceptions of middle and high school students of Mexican descent: A Q methodology study. *Operant Subjectivity*, 33, 103-128.
- Tashakkori, A., & Teddlie, C. (2010). *Sage handbook of mixed methods in social & behavioral research*. Thousand Oaks, CA: Sage.
- Taylor, A., Harold, J., & Berger, K. (2013, June 17). The overhead myth. Retrieved from <http://overheadmyth.com/letter-to-the-donors-of-america/>
- Thomas, D. B., & Baas, L. R. (1992). The issue of generalization in Q methodology: "Reliable schematics" revisited. *Operant Subjectivity*, 16, 18-36. Retrieved from <http://operantsubjectivity.org/os/about>
- Thomson, D. E. (2010). Exploring the role of funders' performance reporting mandates in nonprofit performance measurement. *Nonprofit and Voluntary Sector Quarterly*, 20(10), 1-19. doi:10.1177/0899764009360575
- Tierney, T. J., & Steele, R. (2011). *The donor-grantee trap: How ineffective collaboration undermines philanthropic results for society, and what can be done about it*. Boston, MA: The Bridgespan Group.
- Tinkleman, D., & Donabedian, B. (2007). Street lamps, alleys, ratio analysis, and nonprofit organizations. *Nonprofit Management & Leadership*, 18, 5-19. Retrieved from <http://onlinelibrary.wiley.com/journal/10.1002/%28ISSN%291542-7854>
- Toepler, S. (2003). Grassroots associations versus larger nonprofits: New evidence from a community case study in arts and culture. *Nonprofit and Voluntary Sector Quarterly*, 32, 236-251. doi:10.1177/0899764003032002004

- Twersky, F., Buchanan, P., & Threlfall, V. (2013). Listening to those who matter most, the beneficiaries. *Stanford Social Innovation Review*, 11(2), 40-45.
- Unerman, J., & O'Dwyer, B. (2006). Theorising accountability for NGO advocacy. *Accounting, Auditing & Accountability Journal*, 19, 349-376. Retrieved from <http://www.emeraldgroupublishing.com/products/journals/journals.htm?id=AAAJ>
- Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy. (2004a). *Getting what we pay for: Low overhead limits nonprofit effectiveness*. Retrieved from <http://nccsdataweb.urban.org/kbfiles/311/brief%203.pdf>
- Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy. (2004b). *The pros and cons of financial efficiency standards*. Retrieved from <http://nccsdataweb.urban.org/kbfiles/521/brief%205.pdf>
- Urban Institute Center on Nonprofits and Philanthropy & Indiana University Center on Philanthropy. (2004c). *Who raises contributions for America's nonprofit organizations?* Retrieved from <http://nccsdataweb.urban.org/kbfiles/312/Brief%202.pdf>
- U.S. General Accounting Office. (2002). *Tax-exempt organizations: Improvements possible in public, IRS, and state oversight of charities* (Report to the Chairman and Ranking Minority Member, Committee on Finance, U.S. Senate No. GAO-02-526). Retrieved from <http://www.gao.gov/new.items/d02526.pdf>
- U.S. Government Accountability Office. (2007). Nonprofit sector: Increasing numbers and key role in delivering federal services. (Testimony before the Subcommittee

- on Oversight, Committee on Ways and Means, U.S. House of Representatives No. GAO-07-1084T). Retrieved from [www.gao.gov/cgi-bin/getrpt?GAO-07-1084T](http://www.gao.gov/cgi-bin/getrpt?GAO-07-1084T)
- Valaitis, R., Akhtar-Danesh, N., Eva, K., Levinson, A., & Wainman, B. (2007). Pragmatists, positive communicators, and shy enthusiasts: three viewpoints on web conferencing in health sciences education. *Journal of Medical Internet Research*, 9(5), e39. doi:10.2196/jmir.9.5.e39
- van Exel, J., & de Graaf, G. (2005). *Q methodology: A sneak preview*. Retrieved from <http://www.jobvanexel.nl>
- Watts, S. & Stenner, P. (2012). *Doing Q methodological research: Theory, method, and interpretation*. Newbury Park, CA: Sage.
- Webler, T. D., Danielson, S., & Tuler, S. (2009, February). *Using Q method to reveal social perspectives in environmental research*. Retrieved from <http://www.serius.org/pubs/Qprimer.pdf>
- Wecht, C. H., & Rago, J. T. (2006). *Forensic science and law: Investigative applications in criminal, civil, and family justice*. Boca Raton, FL: CRC Press.
- Wellar, K. (2012, July 30). Embracing risk in the shift from “program thinking” to “social change thinking.” *Management: Nonprofit Quarterly*. Retrieved from <http://nonprofitquarterly.org/>
- Wheatley, M., & Kellner-Rogers, M. (1999, June). What do we measure and why? Questions about the uses of measurement. *Journal of Performance Management*, 24-27.

- Whetten, D. A. (1978). Coping with incompatible expectations: An integrated view of role conflict. *Administrative Science Quarterly*, 23, 254-271. Retrieved from <http://asq.sagepub.com/>
- Whitaker, G. P., Altman-Sauer, L., & Henderson, M. (2004). Mutual accountability between governments and nonprofits. *American Review of Public Administration*, 34, 115-133. Retrieved from <http://arp.sagepub.com/>
- Wing, K. T. (2009). Overhead is overrated in financial reporting. *Chronicle of Philanthropy*, 21(22), 34-37. Retrieved from <http://www.philanthropy.com>
- Wing, K. T., Pollak, T. H., & Blackwood, A. (2008). *The nonprofit almanac 2008*. Washington, DC: The Urban Institute.
- Wolf, A. (2010). Orientations to academic workloads at department level. *Educational Management Administration & Leadership*, 38, 246-264.  
doi:10.1177/1741143209356362
- Worth, M. J. (2009). *Nonprofit management: Principles and practice*. Thousand Oaks, CA: Sage.
- Yuchtman, E., & Seashore, S. E. (1967). A system resource approach to organizational effectiveness. *American Sociological Review*, 32, 891-903. Retrieved from <http://www.asanet.org/journals/asr/>
- Zammuto, R. F. (1982). *Assessing organizational effectiveness: Systems change, adaptation, and strategy*. New York, NY: State University of New York Press.
- Zammuto, R. F. (1984). A comparison of multiple constituency models of organizational effectiveness. *Academy of Management Review*, 9(4), 606-616. Retrieved from <http://www.aom.pace.edu/amr/>

Zimmermann, J. M., & Stevens, B. W. (2006). The use of performance measurement in South Carolina nonprofits. *Nonprofit Management & Leadership*, 16, 315-328.

Retrieved from

<http://onlinelibrary.wiley.com/journal/10.1002/%28ISSN%291542-7854>

Zumdahl, L. (2013, July 10). Real talk about real costs: Overhead. *Nonprofit Knowledge Matters: National Council of Nonprofits*. Retrieved from

<http://www.councilofnonprofits.org/>

## Appendix A

### Participation Request Letter

Participant Name  
Address, City, State, Zip

Dear \_\_\_\_\_,

I am a student working on a Doctor of Management in Leadership and Organizational Theory at the University of Phoenix. I am conducting a research study that may help to identify characteristics of effective nonprofit organizational performance measurement. The purpose of the research is to study the opinions of leaders who have experience in nonprofit management and performance measurement and reveal findings that could help evaluate nonprofit performance across the sector.

More specifically, nonprofit leaders with experience in nonprofit management and performance measurement are targeted for this study. For the purposes of the proposed study, nonprofit leaders will be (a) academic experts who have published within the field in the past 5 years or (b) professionals with 5 or more years of leadership experience within the field. Professionals are NPO senior management, evaluators, grantors and funders, consultants, and watchdog agency leaders. You have been selected to participate in this research program, because I believe that you meet the qualifications necessary for inclusion in this study.

As a leader in the nonprofit sector, your participation will involve completing an interview and assessment that will examine your perceptions of the characteristics of effective nonprofit organizational performance measurement.

The assessment should not take more than an hour of your time. Once you have signed the consent form and returned the form in the stamped, self-addressed envelope, you will receive instructions outlining the next steps in the research process. There are no foreseeable risks to you in undertaking this study.

The intent of the study is to produce relevant data that can be used to create a nonprofit organizational performance measurement method that will not only provide relevant performance data to external parties, but also help nonprofit leaders improve organizational performance. An executive summary of the research results will be provided if requested on the consent form.

#### ***Confidentiality Statement***

Your participation in this study is voluntary. If you choose not to participate or to withdraw from the study at any time, you can do so without penalty or loss of benefit to yourself. The



results of the research study may be published, but your name and your organization's name will not be used, and your results will be maintained in strictest confidence.

If you have any questions concerning the research study, please contact me at ###-###-#### or xxxx@email.phoenix.edu .

Sincerely,

Diana Frayne  
Doctoral Student  
University of Phoenix

## Appendix B

### Informed Consent

#### **INFORMED CONSENT: PARTICIPANTS 18 YEARS OF AGE AND OLDER**

Dear \_\_\_\_\_,

My name is Diana Frayne and I am a student at the University of Phoenix working on a Doctor of Management degree. I am doing a research study entitled Nonprofit Leader Perceptions of Effective Organizational Performance Measurement: A Q methodology Study. The purpose of the research study is to identify commonalities among the subjective perspectives of nonprofit leaders in the United States concerning effective organizational performance measurement practices.

Your participation will involve completing a card sorting process and brief questionnaire that will examine your perceptions of the characteristics of effective nonprofit organizational performance measurement. More specifically, nonprofit leaders with experience in nonprofit management and performance measurement are targeted for this study. For the purposes of the proposed study, nonprofit leaders will be (a) academic experts who have published within the field in the past 5 years or (b) professionals with 5 or more years of leadership experience within the field. Professionals are NPO senior management, evaluators, grantors and funders, consultants, and watchdog agency leaders. You have been selected to participate in this research program, because I believe that you meet the qualifications necessary for inclusion in this study.

The assessment should not take more than an hour of your time. Once you have signed the consent form and returned the form in the stamped, self-addressed envelope, you will receive instructions outlining the next steps in the research process. You can decide to be a part of this study or not. Once you start, you can withdraw from the study at any time without any penalty or loss of benefits. Participation in this study is voluntary. Your decision to participate or not participate will not affect your current or future relations within your organization. If you decide to participate, you are free to withdraw at any time or not to answer any question. The results of the research study may be published but your identity will remain confidential and your name will not be made known to any outside party.

In this research, there are no **foreseeable risks** to you in undertaking this study.

Although there may be no direct benefit to you, a possible benefit from your being part of this study is the identification of relevant data that can be used to create a nonprofit organizational performance measurement method that will not only provide relevant performance data to external parties, but also help nonprofit leaders improve

organizational performance. An executive summary of the research results will be provided if requested below.

If you have any questions about the research study, please call me at ###-###-#### and xxxxx@email.phoenix.edu. For questions about your rights as a study participant, or any concerns or complaints, please contact the University of Phoenix Institutional Review Board via email at IRB@phoenix.edu.

As a participant in this study, you should understand the following:

1. You may decide not to be part of this study or you may want to withdraw from the study at any time. If you want to withdraw, you can do so without any problems.
2. Your identity will be kept **confidential**.
3. Diana Frayne, the researcher, has fully explained the nature of the research study and has answered all of your questions and concerns.
4. If interviews are done, they may be recorded. If they are recorded, you must give permission for the researcher, Diana Frayne, to record the interviews. You understand that the information from the recorded interviews may be transcribed. The researcher will develop a way to code the data to assure that your name is protected.
5. Data will be kept in a secure and locked area. The data will be kept for three years, and then destroyed.
6. The results of this study may be published.

“By signing this form, you agree that you understand the nature of the study, the possible risks to you as a participant, and how your identity will be kept confidential. When you sign this form, this means that you are 18 years old or older and that you give your permission to volunteer as a participant in the study that is described here.”

I accept the above terms.       I do not accept the above terms. (CHECK ONE)

Signature of the interviewee \_\_\_\_\_ Date \_\_\_\_\_

Signature of the researcher \_\_\_\_\_ Date \_\_\_\_\_

Please send me an Executive Summary of Research Results

## Appendix C

### Q sample Statements

- \_\_\_(1) A new performance measurement model can help improve philanthropy by identifying chronic undercapitalization and by shifting the unbalanced power dynamic resulting from NPO resource dependencies.
- \_\_\_(2) A new performance measurement model should ensure all constituencies benefit. It must identify each player that contributes to or uses the performance measurement system and how they gain value from it.
- \_\_\_(3) A new performance measurement model should help the NPO learn and improve performance – to drive internal reflection, improve mission execution, communicate strengths and weaknesses, and enhance adaptability.
- \_\_\_(4) A performance measurement system that simplifies the complexity of NPO performance into single measures will continue to prevent the sector to truly improve performance. Effectiveness is multi-dimensional.
- \_\_\_(5) A simple method for NPO leaders to identify the 3 to 5 most important measures they will use to assess their programs from 3 simple categories: 1) how much do we do? 2) how well do we do it? 3) is anyone better off?
- \_\_\_(6) A substantial flaw of many performance metrics is that a high performing organization, defined by that narrow set of indicators, could still pollute, violate laws, and pillage the commons. Furthermore, when there is community benefit, rarely are NPOs the sole contributor to it.
- \_\_\_(7) An effective performance measurement model should include indicators not only for organizational health and program performance, but also for social and economic performance that give a real sense of the organization's outcomes and progress in meeting its vision.
- \_\_\_(8) An effective performance measurement model will examine a range of measures uniquely defined for each organization or type of NPO, identifying the 'intended impact' – the result the NPO will hold itself accountable for. These measures should be mission driven and tailored to each organization because the vast range of difference in NPO goals makes broad measurement of goal attainment impossible to compare.
- \_\_\_(9) An effective performance measurement model will select measurement indicators tied to the organization's strategic plan and internal goals.

- \_\_\_(10) An effective performance measurement model would place less emphasis on financial measures that when required to be low, effectively suffocate NPOs' infrastructure and reduce performance capabilities.
- \_\_\_(11) At times, it seems as if the effectiveness and performance measurement movement has strayed too far from why we should be concerned about effectiveness at all – the passion for solving the social inequities that inspire philanthropy. The movement must focus on measuring and improving service delivery and the NPO's ability to help customers.
- \_\_\_(12) Culture matters more than metrics. Performance measurement must be seen as helping people do their jobs better, not creating new chains for yanking.
- \_\_\_(13) Devoid in the performance measurement discussion is who will pay for it. Since performance measurement is often an underfunded overhead expense, donors who want performance data should match their expectations with their resource commitments so that NPOs can direct resources toward the mission.
- \_\_\_(14) Donors need comparative data and ratings providing actionable information to help build donation strategies and identify seminal NPOs from each field.
- \_\_\_(15) Donors will have a more complete picture of NPOs when deciding to make charitable gifts if there is a clearly defined metric that allows for NPO performance assessment and comparisons, determines the value created, increases transparency, and holds NPOs accountable.
- \_\_\_(16) Externally driven performance measurement will be counterproductive or ill-fitting; internally driven performance measurement will likely fall short in meeting the external data demands or be poorly executed. Ideal performance measurement will blend internal and external parties, the needs of all stakeholders, and performance monitoring and evaluation methods to achieve balance in resources, knowledge and results.
- \_\_\_(17) Financial efficiency will always be the most important performance measurement since NPOs cannot exist without financial inputs from donors. How a NPO uses its resources should be the only measure we are concerned with as it impacts everything else the organization does.
- \_\_\_(18) If NPOs want to prove legitimacy, they need to uphold shared values, increased transparency, and maintain an evidence-based dialogue with their multiple constituencies.
- \_\_\_(19) In our current system, funders dictate what reports NPOs generate, and NPOs generate reports for each funder. However, NPOs are best positioned to know what information is best for them to track. The information tracked should be

mission critical for the NPO, and it should be the information funders should be most interested in receiving.

- \_\_\_(20) It's not hard to imagine NPO performance measurement metrics that could easily distinguish NPOs as leaders within their respective field. It is much harder, however, to compare them across fields.
- \_\_\_(21) Measuring and using outcomes data is essential for improving productivity. Outcome measurement is essential for a NPO to identify where the line is. Learning organizations don't just care about whether they achieve results, but why and how they can get better.
- \_\_\_(22) Metrics of performance answer: 1) how well the NPO is doing its mission and 2) how you know how well the NPO is doing. However, assessment of results against strategy remains a significant challenge for NPOs: staff struggle to determine the right data and collection procedures, due to technical challenges, inadequate resources and support, and lack of capacity and skill.
- \_\_\_(23) Most NPO leaders don't communicate results or lack the skills/knowledge to conduct effective performance measurement, making external evaluation by an unbiased outside source (such as charity watchdogs, government regulators, or independent nonprofit professionals) in the form of unannounced external performance audits necessary/ideal.
- \_\_\_(24) NPO capacity, or lack thereof, impacts effective performance measurement. A primary objective should be to enhance capacity first so that NPO leaders have the tools to measure performance.
- \_\_\_(25) NPO leaders must commit to self-discipline and self-accountability and create a culture of measurement. With determination and discipline, clearly defined goals, and internal benchmarks, NPO leaders will gain clarity in desired results, how to achieve those results, determine what to measure, learn what works/doesn't, and explicitly apply what they learned to improve.
- \_\_\_(26) NPO leaders see few frameworks that seem relevant to them, few peers who can point the path to victory, and, most disappointingly, few funders willing to pay to build the required internal measurement capacity. Most metrics are cumbersome, ill-fitting, bludgeon smaller NPOs with onerous requirements, and are not in the interest of helping NPOs.
- \_\_\_(27) NPO performance measurement is receiving increased academic and practitioner attention, but the design of performance measurement systems has not received the same level of consideration. The design methodology and tools can dictate success or failure.

- \_\_\_(28) NPO performance measurement metrics must never trump the NPO's ability to achieve its mission. Focus should be on accountabilities that advance mission achievement.
- \_\_\_(29) Outcome assessment is a highly complicated, uncertain and increasingly contentious undertaking. Charities work on difficult, complex, and sometimes intractable problems. Let's not reduce their appetite or ambition for working on the really hard problems in deference to easier problems that are more susceptible to quick impact and simple measurement.
- \_\_\_(30) People are obsessed with measurements; we want to constantly measure everything around us. It has become such an integral part of life we don't pause and ask why we are measuring and what are the metrics we are using. A new performance measurement model must challenge the questions and metrics thrust on the sector by all constituencies.
- \_\_\_(31) Performance data should stand on its own, minimizing ratings, recommendations, and expert analysis. We need a national performance measurement apparatus that can communicate the results.
- \_\_\_(32) Performance measurement methods are forced upon NPO leaders because donors do not trust NPO leaders to manage donations. Increased funding levels raise the burden of proof, forcing leaders to be accountable to those they serve.
- \_\_\_(33) Performance measurement should be at the organizational level of performance and allow for experimentation, the essence of innovation. Each measure should be: objective, easily understood, timely, accurate, cost-effective, useful, motivating, and able to be tracked.
- \_\_\_(34) Performance measurement should begin with the end in mind. The desired end state will inform goals, strategies, milestones/indicators, and what data is necessary to collect to effectively measure performance and continually improve.
- \_\_\_(35) Rather than asking traditional strategic planning questions (i.e. goal identification), performance measurement should begin by asking "who is our constituency and what kind of organization do they need us to be?" and "what does our community need us to be doing right now?" A NPO should be judged by how effective they are at helping people and impacting the community.
- \_\_\_(36) Roughly 3,000 charities garner 60% of all revenue that comes into the sector each year, and 20,000 garner 85%. The initial goal should be to rate the 20,000 commanding the lion's share of financial resources and not burden the smaller NPOs with ill-fitted models incompatible with their current capacity.
- \_\_\_(37) Standardized metrics and reporting systems are unlikely to universally align with more NPOs or be sensitive enough to provide useful lessons.

- \_\_\_(38) The focus of annual performance measurement is accountability, whereas multiple times per year is focused on performance improvement. While a blend of formative and summative assessment is ideal – how can we expect NPOs to achieve their mission while simultaneously conducting frequent performance measurement?
- \_\_\_(39) The sector desperately needs a performance measurement model that is rigorous enough to provide the right signals, affordable enough that NPOs can undertake it in the first place, fast enough so NPOs can make decisions, and understandable enough to transfer skills and build capacity. Benefits would also include better results for beneficiaries, lower costs to learn, rapid innovation, improved resource allocation, and lower program costs.
- \_\_\_(40) The sole purpose of NPO performance measurement is to communicate the achievement of socially positive results – answering to the public, “what positive impact did we make?”
- \_\_\_(41) Too many NPOs are concerned with sustaining operations rather than solving the problems they were funded to end. Just because a NPO develops an ideal management structure doesn’t mean the NPO will produce positive outcomes or social impact. Management structure, systems, processes or capacity should not be performance measures included in a performance measurement model as they place emphasis away from the mission accomplishment.
- \_\_\_(42) We must reject overhead expense ratios as relevant evaluation metrics. And, the enemy of performance measurement is not just overhead but oversimplification. A lack of resources necessarily leads to oversimplification.
- \_\_\_(43) We must reject overhead expense ratios as relevant evaluation metrics. And, the enemy of performance measurement is not just overhead but oversimplification. A lack of resources necessarily leads to oversimplification.



## Appendix D

### Q sorting Instructions

This research study is examining and attempting to determine characteristics, behaviors, and practices of effective nonprofit organizational performance measurement, and more specifically, the beliefs of nonprofit leaders about this topic. The researcher is interested in what you believe is most important or essential to effective organizational performance measurement.

There is a deck of statement cards regarding nonprofit performance measurement provided in the envelope. The question you are to answer is as follows: *Which characteristics, behaviors, and practices will promote effective nonprofit organizational performance measurement?* Begin by reading each statement and placing each statement in one of three piles: (1) important, (2) I am neutral, or (3) unimportant. There is no right or wrong way to sort the card statements; the choices are subjectively yours and are inherently correct in any order.

1. From the IMPORTANT pile, select two cards with which you believe are most important and write those two numbers in the column labeled "+5".
2. From the UNIMPORTANT pile, select two cards with which you believe are most unimportant and write those two numbers in the column labeled "-5".
3. From the IMPORTANT pile, select three more cards with which you believe are most important and write their numbers in the column labeled "+4".
4. From the UNIMPORTANT pile, select three more cards with which you believe are most unimportant, and write their numbers in the column labeled "-4".
5. From the IMPORTANT pile, select four more cards with which you believe are

- most important and write their numbers in the column labeled “+3”.
6. From the UNIMPORTANT pile, select four more cards with which you believe are most unimportant and write their numbers in the column labeled “-3”.
  7. From the IMPORTANT pile, select four more cards that believe are most important with your viewpoint; write their statement numbers in the column labeled “+2”.
  8. If you do not have enough cards in the IMPORTANT pile, select them from the I AM NEUTRAL pile or the UNIMPORTANT pile.
  9. From the UNIMPORTANT pile, select four more cards that you believe are most important; write their statement numbers in the column labeled “-2”.
  10. If you do not have enough cards in the UNIMPORTANT pile, select them from the I AM NEUTRAL pile or the IMPORTANT pile.
  11. Continue the back and forth with the five cards for each of the columns labeled “+1” and “-1”, making sure to record the statement numbers.
  12. If you do not have enough cards to fill the columns, take them from the I AM NEUTRAL pile or the next adjacent pile.
  13. Write down the remaining six card numbers in the column labeled “0” for I AM NEUTRAL.
  14. When you are completed, there should be no cards left over and no blank spaces shown on the answer grid.

Thank you for your time and for participating in the study.

## Appendix E

### Pre- and Postexercise Questionnaire

#### Demographic Survey

Demographic analysis assists in further qualifying the subject population and provides additional information that can be referenced to literature related to nonprofit organizational performance measurement.

As part of the interviewing exercise, please circle or check the answers that most accurately apply to you in each item.

1. Gender:
  - a. Male
  - b. Female
  
2. Age:
  - a. 25-34
  - b. 35-44
  - c. 45-54
  - d. 55-64
  - e. 65-74
  - f. 75+
  
3. Highest Degree Earned:
  - a. Diploma
  - b. Bachelor's
  - c. Master's
  - d. Master's and specializations
  - e. Doctorate
  - f. Other: \_\_\_\_\_
  
4. Years of Experience in Nonprofit Field:
  - a. 5-9
  - b. 10-14
  - c. 15-19
  - d. 20-24
  - e. 25+
  
5. Current *primary* role in nonprofit field (please circle one):
  - a. 501c3 Executive Director/CEO
  - b. 501c3 Senior Management
  - c. Foundation Executive/Grant or Funding Executor
  - d. Academic/Published in Field
  - e. Charity Watchdog Executive

- f. Consultant
  - g. Other: \_\_\_\_\_
6. Number of years in current *primary* role:
- a. 1-2
  - b. 3-5
  - c. 6-9
  - d. 10-14
  - e. 15+
7. Indicate all roles held during your career in the nonprofit field (circle all that apply):
- a. 501c3 Executive Director/CEO
  - b. 501c3 Senior Management
  - c. Foundation Executive/Grant or Funding Executor
  - d. Academic/Published in Field
  - e. Charity Watchdog Executive
  - f. Consultant
  - g. Other: \_\_\_\_\_
8. Current *primary* organization type you work for/with (please circle one):
- a. Animals
  - b. Arts, Culture, Humanities
  - c. Education
  - d. Environment & Animals
  - e. Health
  - f. Human Services
  - g. International/Foreign Affairs
  - h. Public/Societal Benefit
  - i. Mutual/Membership Benefit
  - j. Religious
  - k. Unknown/Unclassified
  - l. Not applicable
9. Current *primary* organization operating budget (please circle one):
- a. Less than \$499,000
  - b. \$500,000-\$999,999
  - c. \$1 million - \$2,499,999
  - d. \$2.5 million - \$4,999,999
  - e. \$5 million - \$9,999,999
  - f. \$10 million - \$24,999,999
  - g. \$25 million - \$49,999,999
  - h. \$50 million or more
10. Current *primary* organization number of full time employees (please circle one):
- a. 1-10

- b. 11-25
- c. 26-50
- d. 51-100
- e. 101-200
- f. 201 or more

11. Current *primary* organization geographic region (please circle one):

- a. Northeast U.S.
- b. Southeast U.S.
- c. North Central U.S.
- d. South Central U.S.
- e. Northwest U.S.
- f. Southwest U.S.

12. In your current *primary* organization, what role does organizational performance measurement play:

- a. Very important
- b. Important
- c. Somewhat important
- d. Somewhat unimportant
- e. Unimportant
- f. Very unimportant

13. Throughout your career in the nonprofit sector, which organization types have you worked for/with (circle all that apply):

- a. Animals
- b. Arts, Culture, Humanities
- c. Education
- d. Environment & Animals
- e. Health
- f. Human Services
- g. International/Foreign Affairs
- h. Public/Societal Benefit
- i. Mutual/Membership Benefit
- j. Religious
- k. Unknown/Unclassified

### Postsort Questions

Postsor feedback helps with further interpretation of your Q sort results. As part of the interviewing exercise after completing the Q sorting exercise, please respond to the following questions.

1. Were your opinions about effective nonprofit organizational performance measurement appropriately represented in this sort? Why or why not?



Appendix F

Participant Demographic Information

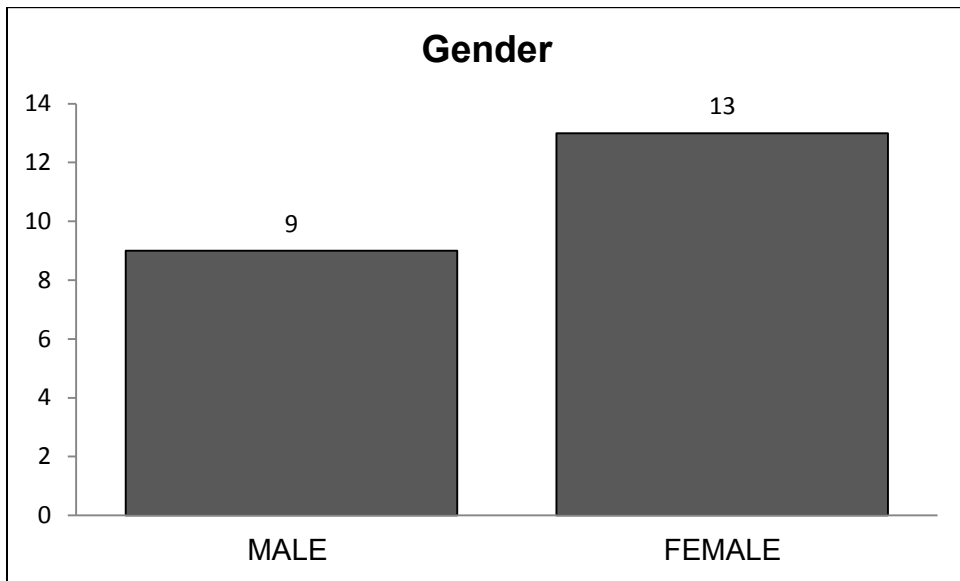


Figure F1. Gender of study participants

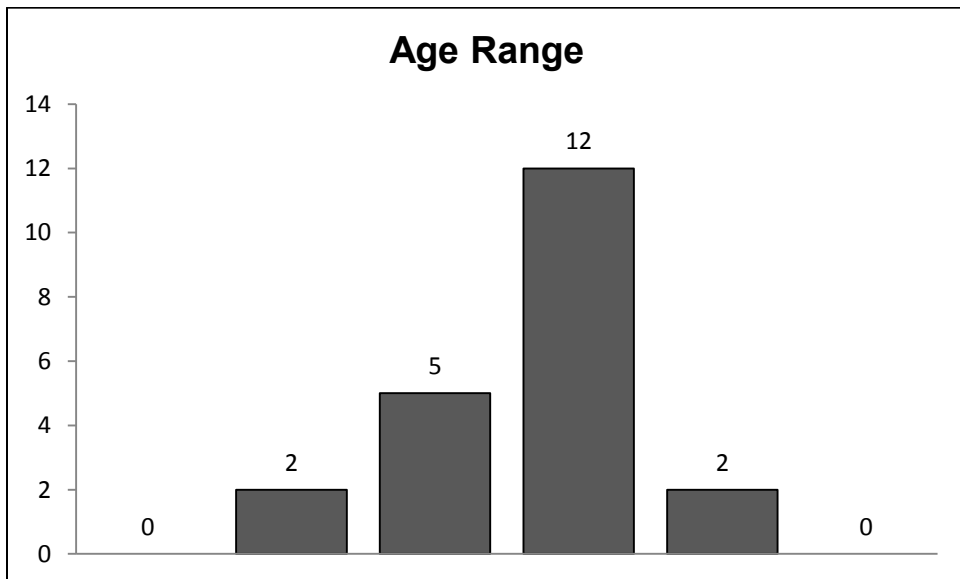
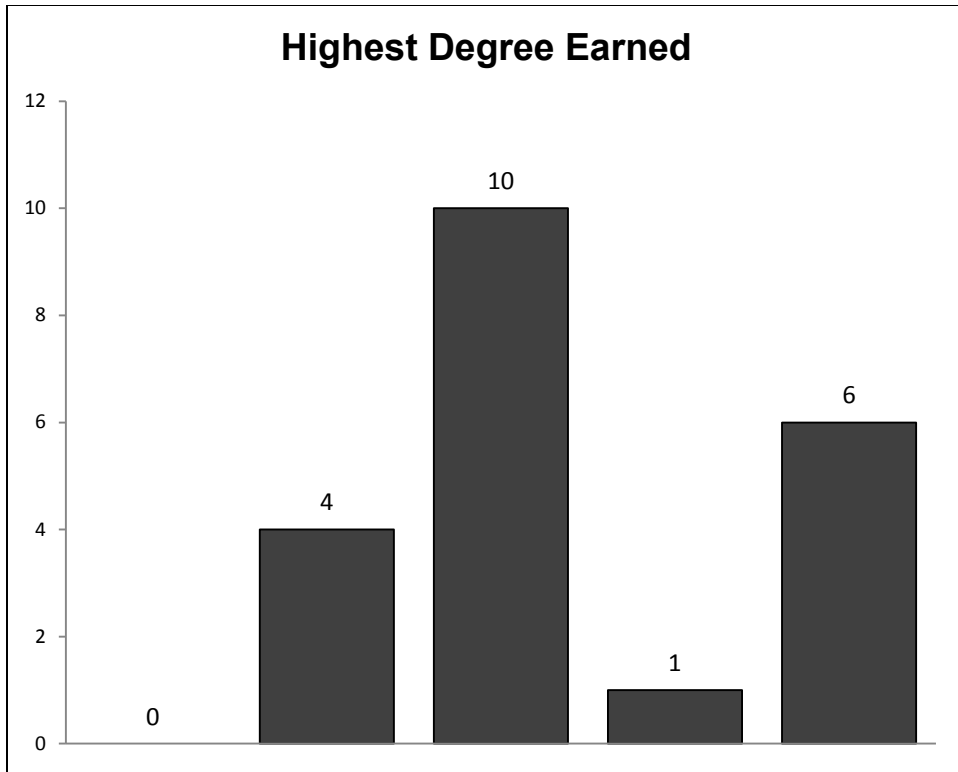
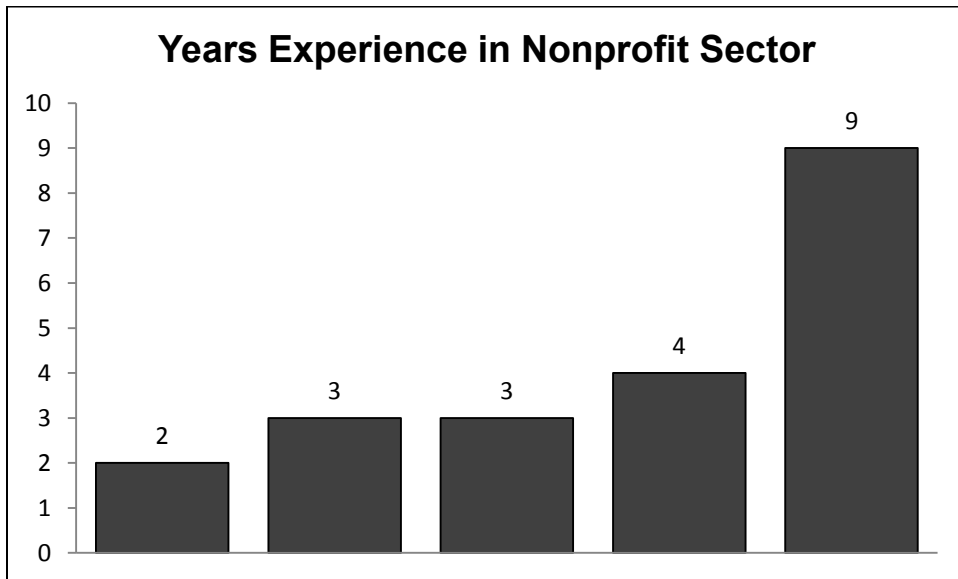


Figure F2. Age ranges of study participants



*Figure F3.* Highest degree earned



*Figure F4.* Number of years experience within the nonprofit sector



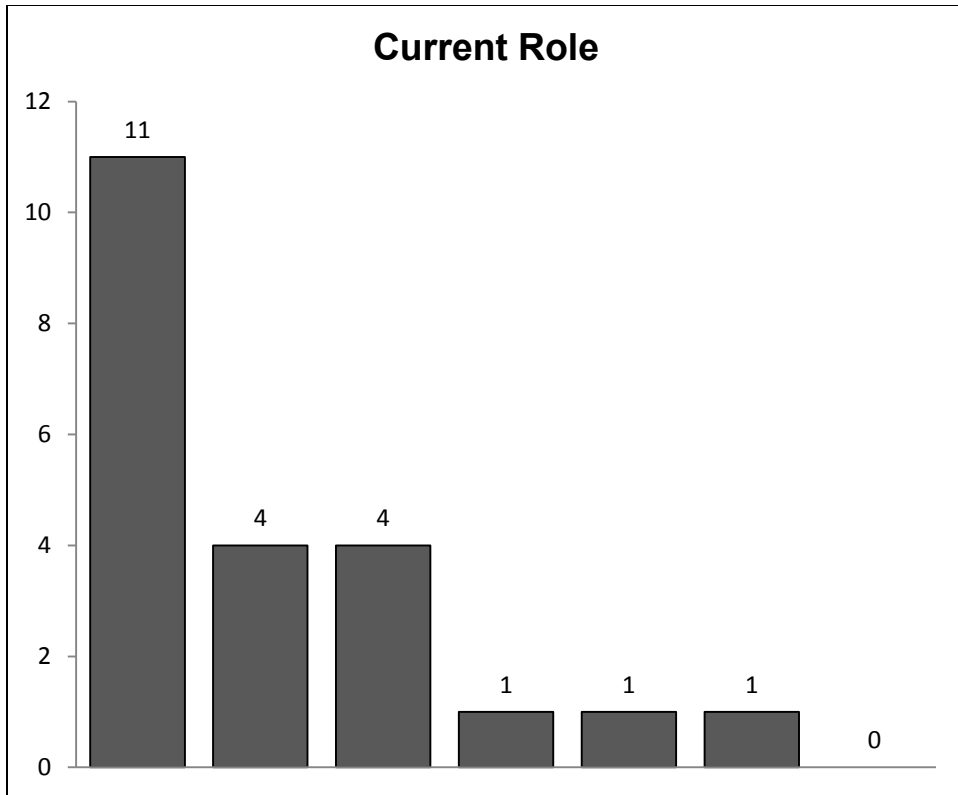


Figure F5. Current role in the nonprofit sector

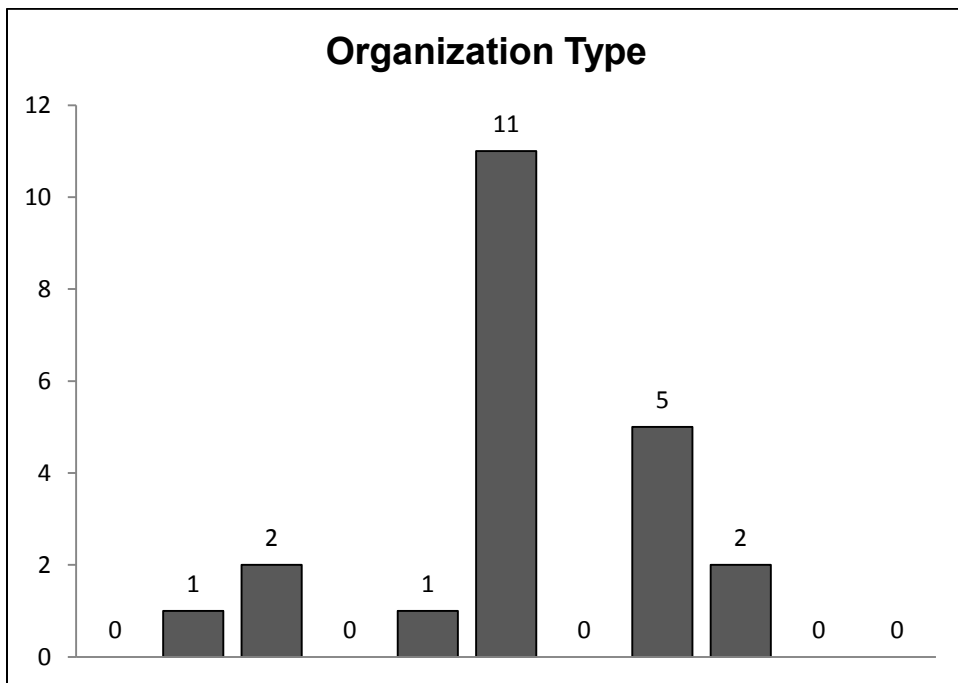


Figure F6. Current organization type of study participants

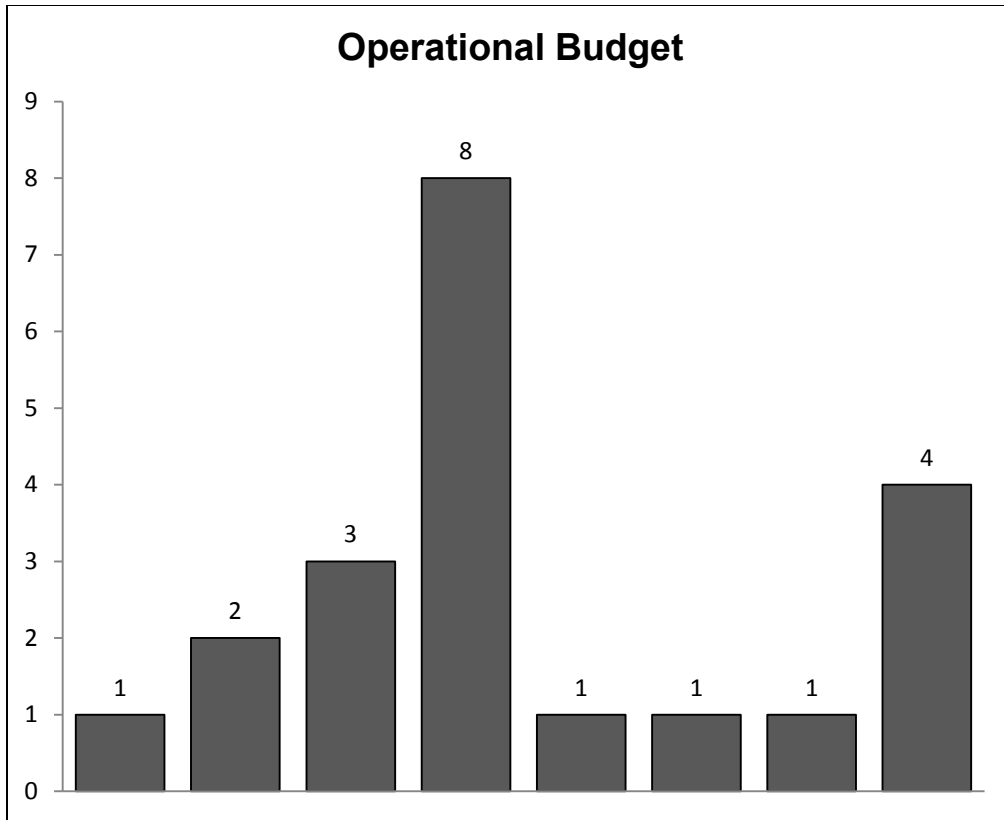


Figure F7. Current organization's operational budget

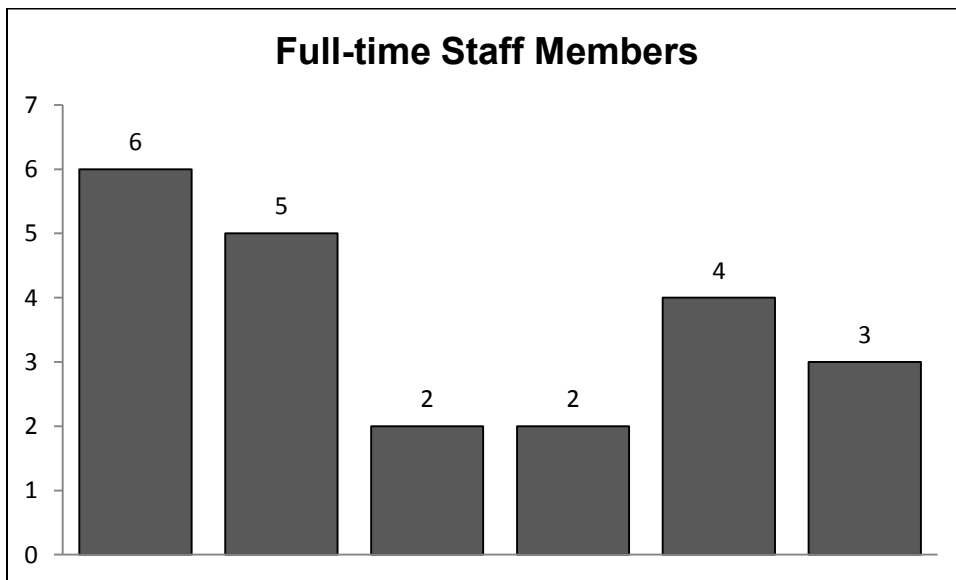


Figure F8. Number of full-time staff members at current organization

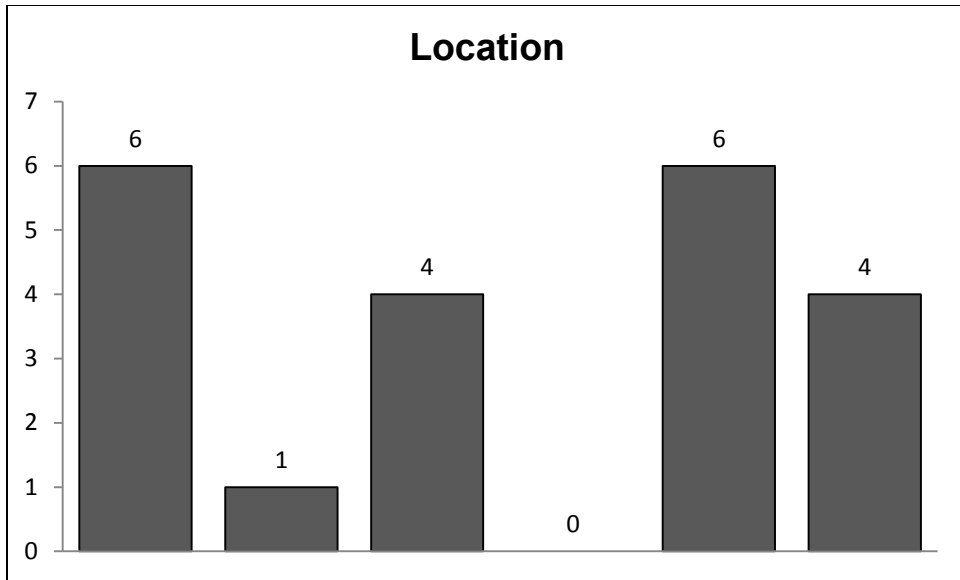


Figure F9. Current organization's location

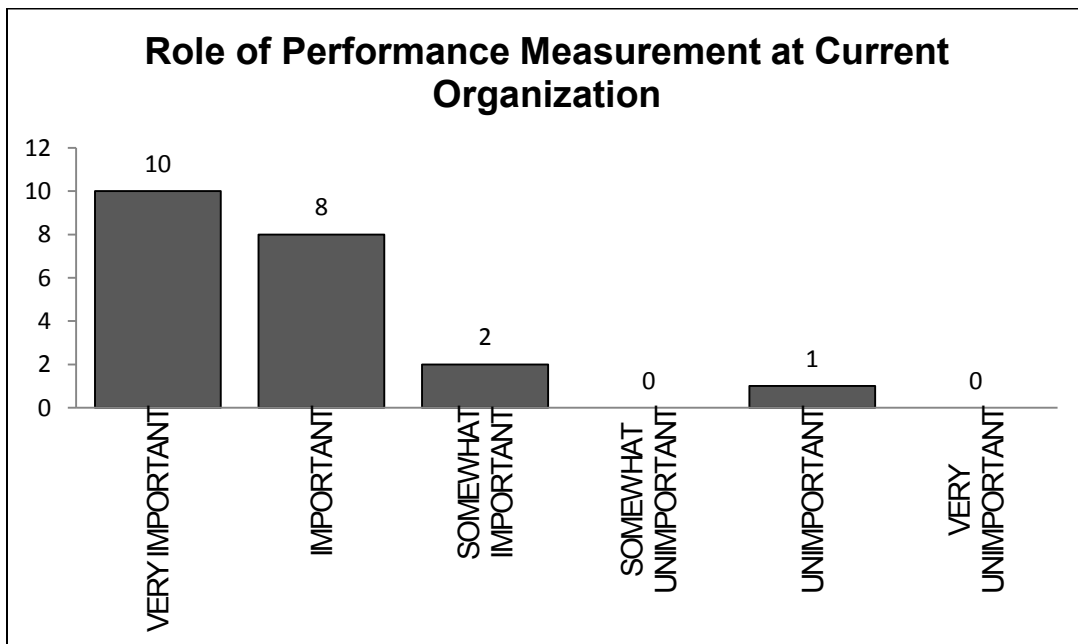


Figure F10. Role of performance measurement at current organization

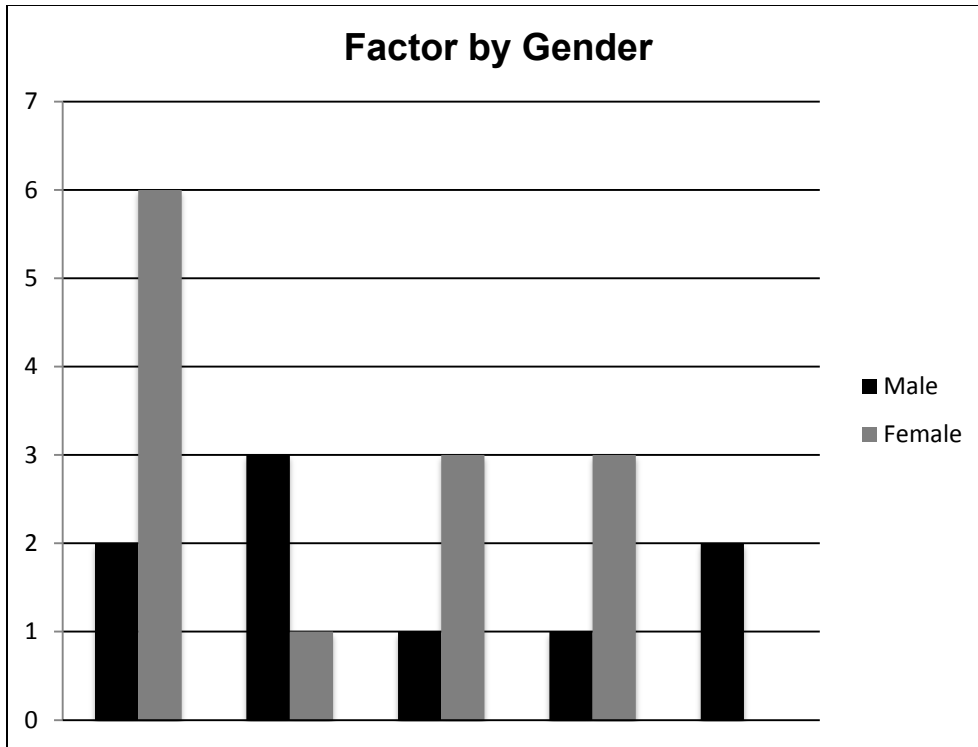


Figure F11. Factor by Gender

## Appendix G

### Full Tables of Most Important/Most Unimportant Statements and Scores by

#### Factor

Table G1

#### *Most Important and Most Unimportant Statements and Scores for Road Map*

Item Number	Statement	Array Position	z-Score
33	Performance measurement should be at the organizational level of performance and allow for experimentation, the essence of innovation. Each measure should be: objective, easily understood, timely, accurate, cost-effective, useful, motivating, and able to be tracked.	5	1.66*
3	A new performance measurement model should help the NPO learn and improve performance – to drive internal reflection, improve mission execution, communicate strengths and weaknesses, and enhance adaptability.	5	1.60
28	NPO performance measurement metrics must never trump the NPO's ability to achieve its mission. Focus should be on accountabilities that advance mission achievement.	4	1.59*
7	An effective performance measurement model should include indicators not only for organizational health and program performance, but also for social and economic performance that give a real sense of the organization's outcomes and progress in meeting its vision.	4	1.53
35	Rather than asking traditional strategic planning questions (i.e. goal identification), performance measurement should begin by asking "who is our constituency and what kind of organization do they need us to be?" and "what does our community need us to be doing right now?" A NPO should be judged by how effective they are at helping people and impacting the community.	4	1.43
9	An effective performance measurement model will select measurement indicators tied to the organization's strategic plan and internal goals.	3	1.10
4	A performance measurement system that simplifies the complexity of NPO performance into single measures will continue to prevent the sector to truly improve performance. Effectiveness is multi-dimensional.	3	0.98*
34	Performance measurement should begin with the end in mind. The desired end state will inform goals, strategies, milestones/indicators, and what data is necessary to collect to effectively measure performance and continually improve.	3	0.89
24	NPO capacity, or lack thereof, impacts effective performance measurement. A primary objective should be to enhance capacity first so that NPO leaders have the tools to measure performance.	3	0.88
15	Donors will have a more complete picture of NPOs when deciding to make charitable gifts if there is a clearly defined metric that allows for NPO performance assessment and comparisons, determines the value created, increases transparency, and holds NPOs accountable.	-3	-0.95
40	The sole purpose of NPO performance measurement is to communicate the achievement of socially positive results – answering to the public, "what positive impact did we make?"	-3	-1.03
41	Too many NPOs are concerned with sustaining operations rather than solving the problems they were funded to end. Just because a NPO develops an ideal management structure doesn't mean the NPO will produce positive outcomes or social impact. Management structure, systems, processes or capacity should not be performance measures included in a performance measurement model as they place emphasis away from the mission accomplishment.	-3	-
36	Roughly 3,000 charities garner 60% of all revenue that comes into the sector each year, and 20,000 garner 85%. The initial goal should be to rate the 20,000 commanding the lion's share of financial resources and not burden the smaller NPOs with ill-fitted models incompatible with their current capacity.	-3	-1.13
14	Donors need comparative data and ratings providing actionable information to help build donation strategies and identify seminal NPOs from each field.	-4	-
23	Most NPO leaders don't communicate results or lack the skills/knowledge to conduct effective performance measurement, making external evaluation by an unbiased outside source (such as charity watchdogs, government regulators, or independent nonprofit professionals) in the form of unannounced external performance audits necessary/ideal.	-4	1.49* -
32	Performance measurement methods are forced upon NPO leaders because donors do not trust NPO leaders to manage donations. Increased funding levels raise the burden of proof, forcing leaders to be accountable to those they serve.	-4	1.60* -1.63
31	Performance data should stand on its own, minimizing ratings, recommendations, and expert analysis. We need a national performance measurement apparatus that can communicate the results.	-5	-2.01
17	Financial efficiency will always be the most important performance measurement since NPOs cannot exist without financial inputs from donors. How a NPO uses its resources should be the only measure we are concerned with as it impacts everything else the organization does.	-5	-2.14

Note: Numbers with asterisks denote distinguishing statements

Table G2

*Most Important and Most Unimportant Statements and Scores for Management Tool*

Item Number	Statement	Array Position	z-Score
8	An effective performance measurement model will examine a range of measures uniquely defined for each organization or type of NPO, identifying the ‘intended impact’ – the result the NPO will hold itself accountable for. These measures should be mission driven and tailored to each organization because the vast range of difference in NPO goals makes broad measurement of goal attainment impossible to compare.	5	1.44*
34	Performance measurement should begin with the end in mind. The desired end state will inform goals, strategies, milestones/indicators, and what data is necessary to collect to effectively measure performance and continually improve.	5	1.42
3	A new performance measurement model should help the NPO learn and improve performance – to drive internal reflection, improve mission execution, communicate strengths and weaknesses, and enhance adaptability.	4	1.39
21	Measuring and using outcomes data is essential for improving productivity. Outcome measurement is essential for a NPO to identify where the line is. Learning organizations don’t just care about whether they achieve results, but why and how they can get better.	4	1.29
18	If NPOs want to prove legitimacy, they need to uphold shared values, increased transparency, and maintain an evidence-based dialogue with their multiple constituencies.	4	1.28
41	Too many NPOs are concerned with sustaining operations rather than solving the problems they were funded to end. Just because a NPO develops an ideal management structure doesn’t mean the NPO will produce positive outcomes or social impact. Management structure, systems, processes or capacity should not be performance measures included in a performance measurement model as they place emphasis away from the mission accomplishment.	3	1.23*
7	An effective performance measurement model should include indicators not only for organizational health and program performance, but also for social and economic performance that give a real sense of the organization’s outcomes and progress in meeting its vision.	3	0.99
25	NPO leaders must commit to self-discipline and self-accountability and create a culture of measurement. With determination and discipline, clearly defined goals, and internal benchmarks, NPO leaders will gain clarity in desired results, how to achieve those results, determine what to measure, learn what works/doesn’t, and explicitly apply what they learned to improve.	3	0.97
13	Devoid in the performance measurement discussion is who will pay for it. Since performance measurement is often an underfunded overhead expense, donors who want performance data should match their expectations with their resource commitments so that NPOs can direct resources toward the mission.	3	0.94
6	A substantial flaw of many performance metrics is that a high performing organization, defined by that narrow set of indicators, could still pollute, violate laws, and pillage the commons. Furthermore, when there is community benefit, rarely are NPOs the sole contributor to it.	-3	-0.90
30	People are obsessed with measurements; we want to constantly measure everything around us. It has become such an integral part of life we don’t pause and ask why we are measuring and what are the metrics we are using. A new performance measurement model must challenge the questions and metrics thrust on the sector by all constituencies.	-3	-0.94*
17	Financial efficiency will always be the most important performance measurement since NPOs cannot exist without financial inputs from donors. How a NPO uses its resources should be the only measure we are concerned with as it impacts everything else the organization does.	-3	-1.07*
4	A performance measurement system that simplifies the complexity of NPO performance into single measures will continue to prevent the sector to truly improve performance. Effectiveness is multi-dimensional.	-3	-1.08
40	The sole purpose of NPO performance measurement is to communicate the achievement of socially positive results – answering to the public, “what positive impact did we make?”	-4	-1.09
42	We must reject overhead expense ratios as relevant evaluation metrics. And, the enemy of performance measurement is not just overhead but oversimplification. A lack of resources necessarily leads to oversimplification.	-4	-1.19
32	Performance measurement methods are forced upon NPO leaders because donors do not trust NPO leaders to manage donations. Increased funding levels raise the burden of proof, forcing leaders to be accountable to those they serve.	-4	-1.51
29	Outcome assessment is a highly complicated, uncertain and increasingly contentious undertaking. Charities work on difficult, complex, and sometimes intractable problems. Let’s not reduce their appetite or ambition for working on the really hard problems in deference to easier problems that are more susceptible to quick impact and simple measurement.	-5	-1.94*
31	Performance data should stand on its own, minimizing ratings, recommendations, and expert analysis. We need a national performance measurement apparatus that can communicate the results.	-5	-2.40

Note: Numbers with asterisks denote distinguishing statements

Table G3

*Most Important and Most Unimportant Statements and Scores for Weakest Link*

Item no.	Statement	Array position	z-score
39	The sector desperately needs a performance measurement model that is rigorous enough to provide the right signals, affordable enough that NPOs can undertake it in the first place, fast enough so NPOs can make decisions, and understandable enough to transfer skills and build capacity. Benefits would also include better results for beneficiaries, lower costs to learn, rapid innovation, improved resource allocation, and lower program costs.	5	1.90*
16	Externally driven performance measurement will be counterproductive or ill-fitting; internally driven performance measurement will likely fall short in meeting the external data demands or be poorly executed. Ideal performance measurement will blend internal and external parties, the needs of all stakeholders, and performance monitoring and evaluation methods to achieve balance in resources, knowledge and results.	5	1.81*
18	If NPOs want to prove legitimacy, they need to uphold shared values, increased transparency, and maintain an evidence-based dialogue with their multiple constituencies.	4	1.45
3	A new performance measurement model should help the NPO learn and improve performance – to drive internal reflection, improve mission execution, communicate strengths and weaknesses, and enhance adaptability.	4	1.40
21	Measuring and using outcomes data is essential for improving productivity. Outcome measurement is essential for a NPO to identify where the line is. Learning organizations don't just care about whether they achieve results, but why and how they can get better.	4	1.27
25	NPO leaders must commit to self-discipline and self-accountability and create a culture of measurement. With determination and discipline, clearly defined goals, and internal benchmarks, NPO leaders will gain clarity in desired results, how to achieve those results, determine what to measure, learn what works/doesn't, and explicitly apply what they learned to improve.	3	1.10
7	An effective performance measurement model should include indicators not only for organizational health and program performance, but also for social and economic performance that give a real sense of the organization's outcomes and progress in meeting its vision.	3	1.09
24	NPO capacity, or lack thereof, impacts effective performance measurement. A primary objective should be to enhance capacity first so that NPO leaders have the tools to measure performance.	3	1.08
35	Rather than asking traditional strategic planning questions (i.e. goal identification), performance measurement should begin by asking "who is our constituency and what kind of organization do they need us to be?" and "what does our community need us to be doing right now?" A NPO should be judged by how effective they are at helping people and impacting the community.	3	1.01
29	Outcome assessment is a highly complicated, uncertain and increasingly contentious undertaking. Charities work on difficult, complex, and sometimes intractable problems. Let's not reduce their appetite or ambition for working on the really hard problems in deference to easier problems that are more susceptible to quick impact and simple measurement.	-3	-0.85*
27	NPO performance measurement is receiving increased academic and practitioner attention, but the design of performance measurement systems has not received the same level of consideration. The design methodology and tools can dictate success or failure.	-3	-0.90*
4	A performance measurement system that simplifies the complexity of NPO performance into single measures will continue to prevent the sector to truly improve performance. Effectiveness is multi-dimensional.	-3	-1.07
26	NPO leaders see few frameworks that seem relevant to them, few peers who can point the path to victory, and, most disappointingly, few funders willing to pay to build the required internal measurement capacity. Most metrics are cumbersome, ill-fitting, bludgeon smaller NPOs with onerous requirements, and are not in the interest of helping NPOs.	-3	-1.26*
36	Roughly 3,000 charities garner 60% of all revenue that comes into the sector each year, and 20,000 garner 85%. The initial goal should be to rate the 20,000 commanding the lion's share of financial resources and not burden the smaller NPOs with ill-fitted models incompatible with their current capacity.	-4	-1.31
37	Standardized metrics and reporting systems are unlikely to universally align with more NPOs or be sensitive enough to provide useful lessons.	-4	-1.47*
32	Performance measurement methods are forced upon NPO leaders because donors do not trust NPO leaders to manage donations. Increased funding levels raise the burden of proof, forcing leaders to be accountable to those they serve.	-4	-1.66
41	Too many NPOs are concerned with sustaining operations rather than solving the problems they were funded to end. Just because a NPO develops an ideal management structure doesn't mean the NPO will produce positive outcomes or social impact. Management structure, systems, processes or capacity should not be performance measures included in a performance measurement model as they place emphasis away from the mission accomplishment.	-5	-1.77*
17	Financial efficiency will always be the most important performance measurement since NPOs cannot exist without financial inputs from donors. How a NPO uses its resources should be the only measure we are concerned with as it impacts everything else the organization does.	-5	-1.87

Note: Numbers with asterisks denote distinguishing statements.