

Summer 8-2018

Staff Members' Perceptions of the Impact of Performance Measurement Requirements on Program Delivery and Organizational Effectiveness in Nonprofit Human Service Agencies

Stacey Tice

Follow this and additional works at: <https://knowledge.library.iup.edu/etd>

Recommended Citation

Tice, Stacey, "Staff Members' Perceptions of the Impact of Performance Measurement Requirements on Program Delivery and Organizational Effectiveness in Nonprofit Human Service Agencies" (2018). *Theses and Dissertations (All)*. 1615.
<https://knowledge.library.iup.edu/etd/1615>

This Dissertation is brought to you for free and open access by Knowledge Repository @ IUP. It has been accepted for inclusion in Theses and Dissertations (All) by an authorized administrator of Knowledge Repository @ IUP. For more information, please contact cclouser@iup.edu, sara.parme@iup.edu.

STAFF MEMBERS' PERCEPTIONS OF THE IMPACT OF PERFORMANCE
MEASUREMENT REQUIREMENTS ON PROGRAM DELIVERY AND
ORGANIZATIONAL EFFECTIVENESS IN NONPROFIT HUMAN SERVICE AGENCIES

A Dissertation

Submitted to the School of Graduate Studies and Research

in Partial Fulfillment of the

Requirements for the Degree

Doctor of Philosophy

Stacey L. Tice

Indiana University of Pennsylvania

August 2018

Indiana University of Pennsylvania
School of Graduate Studies and Research
Department of Sociology

We hereby approve the dissertation of

Stacey L. Tice

Candidate for the degree Doctor of Philosophy

J. Beth Mabry, Ph.D.
Professor of Sociology, Chair

Dana Hysock Witham, Ph.D.
Associate Professor of Sociology

Melissa Swauger, Ph.D.
Associate Professor of Sociology

ACCEPTED

Randy L. Martin, Ph.D.
Dean
School of Graduate Studies and Research

Title: Staff Members' Perceptions of the Impact of Performance Measurement Requirements on Program Delivery and Organizational Effectiveness in Nonprofit Human Service Agencies

Author: Stacey L. Tice

Dissertation Chair: Dr. J. Beth Mabry

Dissertation Committee Members: Dr. Dana Hysock Witham
Dr. Melissa Swauger

This study explores how the implementation of performance measurement affects organizations and staff members of nonprofit human service agencies. Performance measurement is defined in this study as the general activities used to collect and report information about a program's effectiveness and efficiency with the intent of using the information for monitoring, resource allocation, or improvement of the program. Qualitative research methods were used to examine the perceptions nonprofit managers and caseworkers had about the effect of performance measurement and funding accountability requirements on service provision in their organizations. The study included semi-structured interviews of 28 employees, 21 caseworkers and seven managers, across seven agencies who provide employment and training programs.

The results of this study suggest that performance measurement is used primarily as a monitoring tool and not used effectively as a tool for program improvement. Although agencies collect extensive data to demonstrate compliance there is no consideration of factors that affect the reliability and validity of the information concerning client outcomes. More crucially, from the caseworker's perspective the information gathered for performance measurement in the E&T programs did not clearly reflect the goals of those programs.

Not only are the data and data collection processes not very useful for enhancing program delivery, but existing performance measurement activities may interfere with organizational

effectiveness. For example, staff training on data collection and entry into computer systems for reporting have supplanted training for caseworkers on providing quality services to the people served by those programs. Performance measurement activities are intertwined with other program activities, so it is difficult to distinguish the effect of performance measurement activities from other processes. However, it is clear from the perspectives of individuals providing services that feedback received from performance measurement activities is not useful for improving service delivery and by extension improving organizational effectiveness. These findings suggest all stakeholders would benefit from more collaborative processes to design performance measures and processes, and intentional incorporation of ways that performance data will be used to improve programs and organizational effectiveness.

TABLE OF CONTENTS

Chapter		Page
ONE	INTRODUCTION	1
	Recent Socio-Historical Background.....	2
	Context of the Study	5
	Statement of the Problem.....	8
	Significance of the Study.....	9
	Purpose and Questions	12
	Research Approach	13
	Definition of Terms.....	14
	Assumptions.....	15
	Delimitations.....	17
	Chapter Summary	17
TWO	LITERATURE REVIEW	19
	Review of Literature	20
	Emerging Areas of Research	31
	Limitations of Existing Literature.....	33
	Theoretical Framework	34
	Chapter Summary	43
THREE	METHODS	46
	Research Design.....	46
	Research Sample	48
	Data Collection	50
	Ethical Considerations	52
	Data Analysis and Synthesis.....	53
	Issues of Trustworthiness and Limitations	55
	Chapter Summary	57
FOUR	RESULTS	58
	Setting	59
	Findings.....	65
	Summary	93

Chapter	Page
FIVE	DISCUSSION AND CONCLUSIONS96
	The Impact of Performance Measurement on Service Delivery and Organizational Performance.....97
	Training and Time Commitments to Performance Measurement109
	Communicating Performance Measurement Results.....105
	Limitations109
	Direction for Future Research.....110
	Implications for Policy and Practice115
	Conclusions.....119
	REFERENCES122
	APPENDICES137
	Appendix A - Interview Protocol.....137
	Appendix B - Informed Consent Form139
	Appendix C - IRB Approval Letter143
	Appendix D - Coding and Analysis Description145

LIST OF TABLES

Table		Page
1	Agency Descriptions	62
2	Participant Demographics	64
3	Participant Experience and Education	65

CHAPTER ONE

INTRODUCTION

What are the effects of performance measurements on community based nonprofits that provide human services? The demand for accountability that permeates all aspects of human service delivery is a result of the need to document effectiveness and efficiency by funding sources (Blalock, 1999; Hatry, 2013; Martin & Kettner, 1996). In part, this pressure is driven by scandals that have tarnished the public image of nonprofits (Arenson, 1995; LeRoux & Wright, 2010) and public agencies (Courtney, Needell, & Wulczyn, 2004). As nonprofit organizations have increasingly played a role in delivering government services since the 1960's (Corbin, 1999; Hall, 2006), the push to make government more efficient also affects nonprofits receiving government funds (Kim, 2005; Moynihan, et al., 2011). Performance measurement is a strategy widely used to show effectiveness of programs, but it is associated with changes in practices that may change the organizational structures of human service agencies (Benjamin, 2008).

This research intends to provide more information on the attitudes of staff members, both management and caseworkers, through a qualitative study of nonprofit caseworkers and managers about how the requirements of performance measurements affect the quality of program delivery and provision of services within organizations. This study focuses on nonprofit human service agencies that work with anti-poverty programs, specifically Community Action Agencies (CAAs) that provide employment and training programs.

This introductory chapter will first provide the recent social-historic background of performance measurement and the relationship to anti-poverty programs and nonprofit organizations. To further explain the context of the study, I provide general information about nonprofit organizations and Community Action Agencies. This is followed by a statement of the

problem, the purpose and questions studied, and the research approach. I also introduce the significance of the study, and definitions of key terms. The assumptions of the researcher, and delimitations of the study are also included as well as a chapter summary.

Recent Socio-Historical Background

Since the 1980s, federal government agencies are increasingly under pressure to prove that government programs, especially anti-poverty programs, are effective (Dicke & Ott; 1999, Hall, 2003). The current U.S. debt is over 20 trillion dollars (U.S. Department of the Treasury, 2016) and climbing daily. The percentage of all people living in poverty is 15.1% (U.S. Census Bureau 2016), changing little from the 19% of the U.S. population living in poverty in 1964 when Lyndon B. Johnson declared a War on Poverty, or the 14.2% rate of poverty in 1967 when anti-poverty programs were launched (Danziger, Sandefur, & Weinberg, 1994). One of the anti-poverty initiatives launched by Lyndon B. Johnson during his War on Poverty was the creation of Community Action Agencies (CAAs). CAAs were originally authorized through the 1964 Economic Opportunity Act and funded through the Community Services Block Grant (Office of Community Services [OCS], 2018). According the Office of Community Services (OCS, 2018) the stated purpose of the Community Services Block Grant (CSBG) is

To provide assistance to states and local communities, working through a network of Community Action Agencies (CAA) and other neighborhood-based organizations, for the reduction of poverty, the revitalization of low-income communities, and the empowerment of low-income families and individuals in rural and urban areas to become fully self-sufficient. (OCS, 2018, p. 2)

Not everyone viewed the anti-poverty initiatives of the War on Poverty as effective: in 1988 Ronald Reagan said, “we fought a war on poverty and poverty won” (Lehman, 1988, p.37).

It is generally accepted that little progress has been made in the United States in alleviating poverty despite all the programs and the trillions of dollars spent on programs such as Supplemental Nutrition Assistance Program (SNAP), Temporary Assistance to Needy Families (TANF) and Earned Income Tax Credit Programs (EITC) (Meyer & Sullivan, 2012). Part of this perception is due to the words of Ronald Reagan and the policy changes associated with the Supply Side Economics (Abramovitz, 1992). Though it is not the intent of this study to explore the history of social programs, it is important to note that the perception that poverty won expressed by Reagan in 1988, captured the sentiment of many Americans. This marked a shift in social policy that further shifted the burden of family maintenance away from the government and back to individuals (Abramovitz, 1992). The presidential election of 1980 resulted in not only the election of Ronald Reagan but also the decline of Democrat representation through a loss of 33 seats in the United States House of Representatives and the 12 seats in the United States Senate (Hibbs, 1982 p. 387). Those on the political right saw this as the electorate's rejection of the social policies of the War on Poverty and Lyndon Johnson's Great Society (Hibbs, 1982) among other Democratic stances of that era.

Policymakers' responses to concerns about the raising debt amidst the apparent ineffectiveness of government-funded anti-poverty programs have been to adopt policies that incorporate performance measurement. Such measures are intended to inform the allocation of resources to programs that demonstrate effectiveness. In turn this would justify the cost to taxpayers of providing supportive services to people in need by allocating resources to the most effective programs. It also acts to counter the perception that government is losing the War on Poverty and appease the push for smaller government. The incorporation of performance measurement into policies shifted the concept of accountability (Martin & Frahm, 2010).

In the 1960's, accountability focused on financial aspects of organizational performance; today accountability includes the effectiveness of programs and activities themselves (Martin & Frahm, 2010). Since the 1980's, these efforts to increase accountability by documenting effectiveness and efficiency of government funded programs include the New Public Management doctrines of the 1980s, The Government Performance and Results Act (GPRA) of 1993, The President's Management Agenda (PMA) of 2002 and the GPRA Modernization Act of 2010. The results are mixed on whether these measures have had the intended effect on resource allocation at the federal level (Dubnick, 2003; Moynihan, 2006), but the push for outcomes has forced agencies to adopt performance measurement practices and think more like for-profit businesses (Drucker, 2011; Murphy, 2007).

There are two broad reasons that the rise in demand for accountability of social programs at the federal level affects human service nonprofits locally. First, the distinction between nonprofits and public agencies became blurred in the 1990s (Bromley & Meyer, 2017; Grobman, 2008). Before the 1960s the government did not contract human services out to private organizations (Martin & Frahm, 2010). The legislative mechanism that allowed human services to be contracted out to private and public organizations was Title XX of the revised Social Security Act of 1975 and the Comprehensive Employment and Training Act of 1978. The increase in contracting to nonprofit organizations (NPOs) was in part due to the need to accommodate the increase in demand of social services that was created through the War on Poverty (DeHoog, 1985). In the 1980s, under Reagan, government contracting to private organizations significantly expanded to demonstrate the government was getting smaller, becoming more efficient and fiscally responsible, and decentralizing control and involvement (DeHoog, 1985; Van Slyke, 2003). The trends in funding sources for nonprofits are complicated and sometimes research is

conflicting. According to Froelich (1999), private funding declined from 30% in 1980 to 19% in 1996. Grobman (2008) reports that by 1997, 36% of nonprofit revenue came from government sources and only 10% came from private donations with the remaining coming from fees and revenues. Nonprofits are continually looking for ways to become less dependent on outside sources for revenue (Froelich, 1999), but the requirement for accountability continues to permeate the sector.

The second reason for an increased emphasis on accountability for human service agencies is that private sector donors are no longer giving to charities just because they agree with the mission. Scandals in nonprofit organizations, including those involving the United Way, Salvation Army, 9/11 Charities, and more recently the Wounded Warriors Project (Hopkins, 2016) have made the general public and granting organizations wary of giving to charitable nonprofits simply because they have a worthy mission.

As a result of the increased contracting by governments to NPOs and the scandals uncovered, funders have turned to performance measurement and accountability practices to such an extent that these practices have become institutionalized (Hall, 2003). Donors, both foundations and private individuals, expect an increased level of program accountability (Grobman, 2008). United Way, the largest nonprofit organization in the United States (Hendricks, Plantz & Pritchard, 2008), and other leading foundations, including the W. K. Kellogg Foundation, Robert Wood Johnson Foundation, have developed comprehensive evaluation methods to be used by the nonprofits they fund (Hall, 2006).

Context of the Study

Nonprofits have a long history in the United States. They were referred to by De Tocqueville (1835/2007) in 1835 as associations unique to the United States as a way to express

our beliefs and organize around causes. A nonprofit corporation has taken the step of formalizing their structure to accommodate internal revenue requirements and to take advantage of tax-exempt status for their organization. The purpose of a nonprofit corporation is to provide public benefit. For that benefit, a nonprofit is given special legal status. The most commonly associated benefit is that a registered nonprofit does not have to pay taxes on any excess funds or profit. Giving to a nonprofit can provide a tax deduction for the donor. There were 1.44 million nonprofits registered in the United States in 2012 (McKeever & Pettijohn, 2014). A public charity is one of 20 different types of nonprofits. Civic leagues, cemeteries, labor organizations, hospitals, benefit associations, fraternal organizations, credit unions and trusts are just some of the types of nonprofits. They generally fall into one of three types of corporations, a funding agency, a membership organization, or a service agency (Grobman, 2008).

Although there are different types of nonprofits, there are common elements. Nonprofits are required to have a board of directors (BOD). The BOD sets policy for the organization, oversees finances, planning, and monitoring (Grobman, 2008). Nonprofits are also required to have a set of bylaws and if a nonprofit has revenue of more than \$50,000 they are not required to be registered with the IRS, but when nonprofits that are required to file an IRS 990 tax return those forms are made public (Grobman, 2008).

A major distinction between for-profit and nonprofit organizations is that nonprofits are driven by their mission and vision while a for-profit organization seeks to gain a monetary profit by providing a good or service that is bought by another entity (Drucker 2011). The mission statement states the who, what, and why of an organization; it can be anything from feeding the poor to saving the Indiana Bat if it has public benefit. The vision of the organization is what the end goal of these activities will be, eradicate hunger or remove the threat of extinction for a

species. Even if a for-profit company adopts a public service activity, that service is funded through profit.

The relationship between the revenue sources and the client distinguishes a nonprofit from for profit corporations. A nonprofit provides a service that is needed by one agent and paid for by a different agent. For example, food bank gives food to low income individuals that is paid for or donated by other individuals. Endangered species do not have a voice; a nonprofit can raise funds to be their advocate. This makes the concept of customer and client ambiguous. The person paying for the service is not always the same as the person receiving the services. Defining outcomes and success can be dependent on the relationship the stakeholder has to the agency. A client may view success as receiving a month's rent. A funder may view success as reducing the risk of homelessness. In the context of performance measurement, who is defining the successful nature of the program may change the measurement preferred and the whether efforts are perceived as effective. It may also place the organization between two competing forces when establishing their effectiveness (Drucker, 2011).

Many Community Action Agencies are nonprofits that are funded through federal Community Services Block Grant (OCS, 2018). According to the National Community Action Partnership (NACP) (2018) there are over 1000 CAA's across the United States including all 50 states, tribal lands, the District of Columbia, and Puerto Rico that cover 99% of all U.S. counties. Some geographic areas are served through private nonprofit agencies and some areas are served by a Public Community Action Agencies. Regardless of whether a CAA is a public or private organization, it must follow Organizational Standards set forth in the Community Services Block Grant Information Memorandum 138 issued by the Department of Health and Human Services (2015). This memorandum outlines the requirements including governance, fiscal responsibility

and reporting that are required (HHS, 2015). For the purposes of this study the significance between a public and a private CAA are immaterial: the reporting and performance measurement requirements are the same and all CSBG Eligible Entities are required to have representation of low income and elected officials either through their Board of Directors or as an advisory board.

CAAs are unique because every CCA determines their own array of programs based on local community need, resulting in a variety of programs with which the study participants are familiar. At the same time, the legislative requirements for organizational structure and reporting provide a degree of consistency over some of the variety of organizations that exist in human service nonprofits. In Pennsylvania there are 41 such agencies that cover all 67 counties (<http://www.thecaap.org>). CAAs in Pennsylvania provide annual reports to the PA Department of Community and Economic Development (DCED) through the Commonwealth Organization and Planning System (COPOS) (PADCED, 2014). This is a process that was specifically created to respond to legislative mandates set forth in the Government Performance and Results Act of 1993 (United States General Accounting Office, 1997).

While Community Action Agencies are locally driven through their governing bodies to address issues of poverty, they share the same mandated federal requirements. These requirements include performance management through state and federal reporting mechanisms. The private nonprofit CAAs can receive funding from multiple sources as any other nonprofit organization does. This creates a unique network of human service agencies with similar mission and vision working in a variety of ways to address the same problem.

Statement of the Problem

In order to make programs more accountable, some type of performance measurement is necessary (Russ-Eft & Preskill, 2009). Organizations have implemented performance

measurements to comply with funders' accountability requirements and to provide evidence of their effectiveness in achieving the outcomes and goals intended with the funding received. Some researchers argue that organizations need to be flexible and responsive to local and immediate situations if they are to be effective (Argyris, 1976; Senge, 2006). A concern is that performance measurement might stifle the organization's ability to be flexible. There is little research on how the staff members of nonprofit organizations perceive the impact of accountability requirements on their organizations and program delivery to the people they serve. This study explores the views of human services employees on how performance measurement, meant to address organizational accountability, affects the provision of services in nonprofit human service organizations from the perception of staff members.

Significance of the Study

Prior to developing this research, an employment counselor whom I know expressed a sentiment that showed me the importance of the relationship between performance measurements and nonprofit human service delivery. She said that as an employment counselor, she no longer felt responsible for the outcomes of her program. She had no power to determine what clients should do since activities are pre-determined by policy guidelines. In addition, criteria used to judge her performance had no relationship to the quality of the work she did with her clients. (personal communication, March 1, 2012). This counselor's statement expresses the disconnect she felt from the policy makers who determined the type of activity and goals for her clients without knowing the individuals she served. If her experience and outlook is similar to those of a large portion of caseworkers who work directly with clients, the processes established to show accountability and verify effectiveness actually may be detrimental to the achievement of the intended outcomes. Indeed, Dicke (2002) found that human services staff members may report

activities in performance measurements from lists of acceptable activities even though what they reported was not what actually occurred; this is done to comply with accountability requirements. In other words, efforts to capture organizational effectiveness and ways to enhance organizational performance may be based on information that is collected to meet performance measurement requirements but that information does not accurately reflect the activities that are occurring during service delivery.

Performance measurement is a strategy that can provide information on whether an organization is effective, but organizational effectiveness is an ambiguous concept with varied meanings (Pandey et al. 2006; Quinn & Rohrbaugh, 1983). Caseworkers are essential in making nonprofit organizations effective in improving the lives of the individuals they serve. If performance measurement activities are making organizations less able to meet their goals, then funders and organizational leaders need to understand how and why that happens. Although one of the top three management concerns in nonprofits, according to a study by Harvard Business School, is performance measurement (Kaplan, 2001), there is little research from the perspective of the nonprofit staff, either the managers or the caseworkers (Carnochan, Samples, Myers, & Austin, 2014) about how performance measurement affects their work, for good or ill. Understanding the perspective of caseworkers can help inform policy makers about how to improve the process of performance measurement and reduce unintended impact on organizational effectiveness. Research provides some answers on why and how nonprofit organizations conduct performance measurement (Behn, 2003, Benjamin, 2012; Carlson, Kelley & Smith, 2010), but lacks insight on how staff members who provide services directly to the clients perceive the effect of accountability practices on their organization.

Another reason it is important to understand the perspectives of the caseworkers on performance measurement in nonprofit organizations (NPOs) is the ability to make the connection between the effective provision of the service and the fulfillment of clients' needs. Typically, there is no direct connection in NPOs between the product (human services) and resources. Donor stakeholders might provide the resources for many NPOs, but know little about the management or outcomes of NPOs. Demands made of NPOs relative to mission and management may not always be in the best interest of society, clients, or the organizations. On the other hand, clients receiving the services often have minimal stakeholder voice. As McHargue (2003) writes, "This disconnect between the funding and outcome can convolute accountability, performance, and mission" (p. 197).

The disconnection identified by McHargue (2003) also leads to another issue that researchers need to address: whether there is confusion between performance measurement for program evaluation and improvement versus performance measurement merely for compliance with funder accountability reporting requirements (Carnochan et al., 2014; LeRoux & Wright, 2010; Thomson, 2010). What agencies do to meet performance measurement may not meet the standards for program evaluation that researchers would demand (Russ-Eft & Preskill, 2009). The claim that performance measurement is used for program improvement is also tenuous. One important piece necessary for using performance measurement for program improvement is a collaborative approach (Russ-Eft & Preskill, 2009). This approach requires that persons with firsthand knowledge have a voice in the process of evaluating a program (Russ-Eft & Preskill, 2009). This study may provide insight on how and to what extent communication occurs from the perspective of the caseworker. A conference call illustrates the value of this point. The National Association for Community Action Programs (NASCAP) conferees were discussing the new data

collection that began in 2018. One person asked who was going to analyze of all of the data agencies collected. The answer, essentially, was that no one was identified by NASCAP for this task yet, but they hoped it would be analyzed. So, in effect, there are important requirements in place, much data will be collected on programs at a national level, but no plan existed to analyze that data or assist providers in using that data for program improvement.

Purpose and Questions

The purpose of this research is to understand how performance measurement affects the staff of nonprofit human service organizations in their work, service delivery, and program improvement efforts. The intent of the push for performance measures, particularly for accountability, is understandable but researchers are questioning the effect this demand has on nonprofits and their provision of human services (Murphy, 2007; Ebrahim, 2003; Hall, 2003; Benjamin, 2012). One question concerns the diversion of organizational resources into performance management data collection and reporting instead of program delivery (Halachmi, 2005). Is the time spent collecting data interfering with service provision (Benjamin, 2012)? Another question concerns the capacity of the persons assigned to data collection and analyses. This task requires training and commitment (Russ-Eft, & Preskill, 2009). Do the workers feel they have adequate training? Researchers also question whether data collection is used to improve programs or if it is used for accountability and compliance (Dicke & Ott, 1999; Ebrahim, 2003; Murphy, 2007). A required component of using performance measurement for program improvement is communication, or a feedback loop (Russ-Eft, & Preskill, 2009). To what extent does this communication involve the caseworkers? This study addresses how human service staff members perceptions about the implementation of performance measures affects the way they do

their jobs and provide human services. Research questions of this study focus on these areas of concern:

1. What is the impact of performance measurement on service delivery and organizational performance relative to the mission?
2. Do NPO staff members feel they have adequate (a) training and (b) time to collect quality performance measurement data?
3. How are performance measurement results communicated to caseworkers for improving the delivery of programs and services?

Research Approach

This is a qualitative research study using semi-structured interviews of caseworkers (staff members who provide services directly to the client), program managers, and agency administrators working in nonprofit human services. Organizations were selected from among the Community Action Agencies (CAA) in Pennsylvania (PA) that provide employment and training programs. The rationale for this pool of organizations was to recruit study participants that experienced both the uniqueness of CAAs and the shared performance measurement processes they use (<http://www.thecaap.org>). This study focuses on the employment and training programs because they typically do not have a professional staff certification requirement, yet they do have program requirements for implementation and they share specific outcome expectations. Many CAAs implement the Work Ready Program funded through the PA Department of Health and Human Services. According to the Pennsylvania Department of Human Services (2018), the Work Ready program is one of several employment services that provides case management to individuals enrolled in Temporary Assistance for Needy Families (TANF). The purpose of the program is to help participants become ready for other employment programs or opportunities.

This program also prepares individuals to participate in other employment programs such as Employment Advancement and Retention Network (EARN). By focusing on CCAs that also provide Work Ready, I have the opportunity to compare perceptions of staff members delivering similar programs through similar reporting systems, allowing for variations in how data collection occurs within their own agency. This has allowed me to explore staff perceptions of performance measurement in different agencies at the same time keeping some consistency between the agencies.

Definitions of Terms

Performance measurement is a term that has been used in the United States for more than 20 years to address monitoring and accountability in both nonprofit and government sector organizations (Russ-Eft & Preskill, 2009). Terms used to address accountability practices such as performance measurement, and performance management and evaluation, are used throughout scholarly and professional literature on the effectiveness of human services with overlapping meaning and intention. Implicit in the use of these terms are concepts such as outputs, outcomes, impact, and logic models (Carnochan et al., 2014; Hatry, 2002). Performance measurement is defined by Martin and Kettner (1997) as “the regular collection and reporting of information about efficiency, quality, and effectiveness of human service programs” (p.19). For the purposes of this research, performance measurement is defined as the general activities used to collect and report information about a program’s effectiveness and efficiency with the intent of using the information for monitoring, resource allocation, or improvement of the program.

Organizational effectiveness is defined differently by researchers (and consultants) resulting in much debate about its meaning and measurement. As a construct, it is a representation of many different values that are dependent on the literature, stakeholder, or researcher making an

agreed upon definition elusive (Quinn & Rohrbaugh, 1983). Some researchers even question if this concept can even be a useful research variable (Pandy et al. 2007; Steers, 1975).

Organizational effectiveness can be defined as simply fiscal solvency or complexly as the ability to be flexible in decision-making and planning (Argyris, 1976). For nonprofits, one definition of effectiveness is the ability to achieve the mission (Grobman, 2008; Drucker, 2011).

Even this straightforward definition has many components such as public perception, flexibility, and ability to raise enough funds to meet the mission goals. Selznick (1996) posits that maintenance needs are better suited to understanding organizational actions than mission (Tompkins, 2005). For the purposes of this research, the definition of organizational effectiveness is the ability of the nonprofit to achieve its mission. It is important to note that the terms organizational effectiveness and organizational accountability can have overlapping meanings. In the 1960's, accountability focused on financial aspects of organizational performance; today accountability includes the effectiveness of programs and activities themselves (Martin & Frahm, 2010). Part of being accountable to funders and the community is showing that the organization is effective in meeting its goals.

Assumptions

I work in upper level management in the nonprofit human service field. Because of my experience with performance measurement, I have concerns about the disconnect between the information agencies are required to report and the actual experience of caseworkers working with low income individuals. Part of my frustration, as I have seen it, is due to the difficulty in getting the information because of the time it takes and the barriers it places on the staff and the individuals seeking services. Often the data collected does not reflect the actual work of helping and important outcomes are not captured in the performance measurements required. I recognize

that this may not be the case in all organizations or with all programs. My three basic assumptions are as follows:

1. All nonprofits receiving funds from the government or foundations employ some type of performance measurement even if they do not specifically call it performance measurement.
2. Human service nonprofit organizations have some type of worker who interacts directly with clients, though the scope and intensity of this work may vary greatly. In this study, I refer to staff members that work directly with clients as caseworkers.
3. There are different perspectives concerning performance measurement at different levels of the organization: staff members who have direct contact with clients (caseworkers) and managers who are responsible for compliance and reporting likely hold different views of performance measurement that stem from their different responsibilities within their organizations.

Removal of the decision-making process, through rules and performance measures, or the control of activities that are driven by performance measures, takes away the component of accountability away from the caseworker. Tetlock (1985) asserts that research on judgment and accountability that does not consider context, does not tell the complete story. In the same manner, looking at performance measurement only from the perspective of funders and managers misses the effect of performance measures on the interactions between clients and caseworkers. In nonprofit human service organizations, if the intent is to change the circumstances of the client, it is only logical that the interaction between caseworker and client is important to creating that change. Organizational effectiveness must then be affected by the quality of that interaction. How performance measurement affects that interaction between the caseworker and the client is

important to understanding the impact of performance measurement on organizational effectiveness.

Delimitations

As with any qualitative study of this nature, the goal is to come to a deeper understanding of the lived experience of the study participants. In this study, I am exploring the experience of human service workers with implementing performance measurement. There is no expectation of being able to generalize the findings of this study beyond identifying patterns and trends within this pool of study participants. Rather, the aim is to fill a gap in the literature and add to our understanding of the impact of performance measurement on how human services workers and organizations serve clients.

This study also is delimited to employment and training programs specific to Community Action Agencies in Pennsylvania. This is a narrow field of human services and the structure of these organizations may limit how this information can be generalized to other nonprofits providing similar services. At the same time, the commonalities of CCAs allow for examining trends and patterns that can be seen within the context of organizations that have similar reporting requirements and performance measurement histories.

Chapter Summary

Nonprofit organizations play an important role in providing human services to individuals and families. Government grants or contracts and private foundations fund many of these agencies and their services; these funders want to see accountability in nonprofit organizations' use of their funds. Part of that accountability is showing that the services are effective, and by extension the organization is effective. Showing accountability through performance measurement is an accepted practice even though researchers know very little about how they

affect the overall effectiveness of organizations. Although there are significant differences in how agencies are required to implement performance measurements, understanding the perception of staff and professionals may help researchers and administrators use them in meaningful ways to improve programs and service delivery. This research studies performance measurement in a specific sector of the nonprofit industry, Community Action Agencies with employment and training programs, to more fully understand how accountability through performance measurements affects service delivery.

CHAPTER TWO

LITERATURE REVIEW

The focus of this research is understanding how professionals and caseworkers of nonprofit organizations perceive the impact of performance measurement on their work and their agency's human service program delivery, specifically in the employment and training programs administered by Community Action Agencies (CAA) in Pennsylvania (PA). This literature review explores the different aspects of organizations and performance measurement that relate to the push for more accountability in human service organizations.

This literature review addresses the current literature on performance measurement within the nonprofit field. Specific topics discussed are organizational effectiveness, control and conformity, performance measurement for improvement, and organizational learning and feedback. Because this study is in the human services field and is concerned with the impact on staff members within this field; a brief discussion of the literature on human service as a gendered labor is also included. Discussion of the emerging areas in research on performance measurement, as well as limitations of existing literature, are also included.

The theoretical framework for this study is largely based on structural functionalism. In the theoretical framework portion of this chapter I will discuss the relevant topics of Weber's bureaucracy, Merton's latent and manifest functions, and institutional theory. The theoretical framework will include how Merton's work with traditional strain theory informs current views on individual adaptation to stressful environments. Also included is a discussion on how organizational theory views the impact on bureaucracy as it relates to gendered organizations.

Review of Literature

In order to synthesize the literature on performance measurement from the human services nonprofit manager perspective, performance measurement is viewed through three lenses used by researchers. Researchers view performance measurement, as part of the larger construct of organizational effectiveness (Crittenden, 2000; Mitchell, 2012), as part of the control mechanism used for monitoring and resource allocation (Brown, Andersson, & Jo, 2015; Crittendon, 2000), or as an evaluation tool used to understand and improve program effectiveness (LeRoux & Wright, 2010; Macindoe & Barman, 2012). For nonprofits, that includes all stakeholders both clients and funders (Herman & Renz, 1997). The control aspect of performance measurement leads into the administration and public management field that deals with contract compliance and resource allocation (Melkers, 2006; Sanger, 2005). The concept of performance measurement as a tool for improvement brings the evaluation field into this discussion (Carmen & Fredricks, 2010; Liket, Rey-Garcia, & Maas, 2014).

Organizational Effectiveness

Organizational effectiveness is a central theme in organizational literature (Quinn & Rohrbaugh, 1983) and continues to garner debate as to its definition and its measurability (Mitchell, 2012; Pandey, Coursey & Moynihan, 2006; Sanger, Sowa, Selden, & Sandfort, 2004; Steers, 1975). Elements identified as important to organizational effectiveness is varied and an ongoing part of the debate on defining organizational effectiveness. The elements identified as part of this review include human, financial, and social capital (Brown, et al., 2015); goal orientation and fiscal health (Crittendon, 2000; Mitchell, 2012); resources and processes (Herman & Renz, 1997); and outcome measurement (Mitchell, 2012; Packard, 2009). Outcome measurement is synonymous with performance measurement as it is defined for the purposes of

this study, although some researchers look at performance measurement on an organizational level (Brown, et al., 2015; Crittenden, 2000; Packard, 2009) and some on a program level (Carnochan, et al. 2014; Mitchell, 2012).

One theme in the limited research that includes managers of human service nonprofits is organizational effectiveness, specifically how managers perceive it and what is important to having an effective organization. This is where the construct performance measurement is identified. In this usage, performance measurement applies to both effectiveness and capacity (Brown, et al., 2015; Carnochan, et al. 2014; Crittenden, 2000; Mitchell, 2012, Packard, 2009; Thompson, 2010). Organizational effectiveness uses performance measurement as the tool to measure the ability of organizations to meet their goal or mission. *Organizational capacity* also looks at the capacity of an organization to implement quality performance measurement (Brown, et al., 2015; Carnochan, et al. 2014).

The findings are mixed as to how nonprofit managers view performance measurement as important to organizational effectiveness. Mitchell (2012) found that 82% of nonprofit leaders who provide human services internationally viewed outcomes as the predominant way to determine organizational effectiveness. Specifically stated as “outcome accountability” (Mitchell, 2012), the meaning of outcomes accountability is to achieve the goals articulated in the organizational mission. For some leaders this also meant outcome evaluation (Mitchell, 2012). Both outcome accountability and outcome evaluation fall into the category of performance measurement as defined for this paper. Herman and Renz (1997) found that inputs and processes such as a strategic plan and client satisfaction were more important than outcomes. Following procedures was less important than “doing things right” (Herman & Renz, 1997, p. 195). In this study, agency managers stated that they were aware that performance information was presented

differently to different stakeholders and was prone to bias (Herman & Renz, 1997). In accurate reporting was also found by Dicke (2002) who studied a state disabilities agency and its residential facilities providers. She found that staff members in 40% of provider agencies reported intentionally inaccurate reporting by nonprofit managers (Dicke, 2002).

The work of Packard (2009) also supports a managerial view that performance measurement is not a good indicator of organizational effectiveness. Packard (2009) surveyed staff, management, and funders. There were 31 organizational criteria in total for organization performance. Evaluating programs for efficiency and effectiveness through outcome measures ranked in the bottom ten in importance (Packard, 2009). Packard's (2009) study included both management and staff. Managers were more concerned about performance measurement than other staff; this was primarily because of the need to meet funder requirements.

In all three of the studies noted above, participants were different. Packard (2009) surveyed nonprofit staff and managers in work force development and high-risk youth human services. In Herman and Renz's (1997) study, managers from both the funder and provider agencies, were surveyed and interviewed in both health and welfare, and disabilities nonprofits. Mitchell (2012) interviewed leaders of nonprofits in only international human service agencies based in the United States. This variety in types of NPOs demonstrates the difficulty in synthesizing the literature on performance measurement.

Two studies focused more on the capacity issues involved in organizational effectiveness (Brown et al., 2015; Carnochan, et al., 2014). Challenges to obtaining data needed for performance measurement was identified (Carnochan, et al., 2014) including tension between the funder and the staff views on how to capture client progress. The difference between managers' views from those of caseworkers was not always clearly articulated, but management's need for

more aggregate data did not always capture client progress as defined by case workers. This created tension between the managers and caseworkers (Carnochan et al., 2014). Although executives felt that high performance was important to organizational capacity and in turn, organizational effectiveness, human capital was rated as a more important factor than performance measurement (Brown, et al., 2015). This tension between managers and caseworkers coupled with the value of human capital to organizational effectiveness leads to the issue of conformity and control.

Control and Conformity

“Ensuring accountability in human services contracting is a fundamental responsibility of government agencies that enter into contractual agreements with private, nonprofit human services providers” (Dicke, 2002, p. 455). Performance measurement for control and monitoring can be used either internally (LeRoux & Wright, 2010; Packard, 2009), as a response to external pressures (Dehoog, 1985; MacIndoe & Barman 2012; Thompson, 2010) or to address both internal and external demands (Christensen & Ebrahim, 2006; Mitchell, 2012). The response to external pressures is primarily a result of funding requirements, but other stakeholders such as clients and local communities, also play a role (Mitchell, 2012).

Funders can use multiple strategies when developing systems to ensure control and conformity of their subcontractors. Dicke and Ott (1999) identified 10 ways that administrators in government ensure accountability when contracting for human services. Keeping in mind that the reasons that governments at all levels are contracting out human services is twofold: a) reduced costs (Dubnick, 2003) and b) the ability of nongovernment organization to be more flexible and innovative (Kramer, 1994). The 10 methods found by Dicke and Ott (1999) are auditing, monitoring, licensure, markets, contracts, registries, courts, whistle blowing, codes of ethics, and

outcomes-based assessments (OBA). OBA, as defined by Dicke and Ott (1999), is the use of evaluation of the results of a program to determine accountability as opposed to evaluating the inputs and procedures. In this model, monitoring is considered to be ongoing inspection of inputs and process as well as outcomes. Contracting is used to determine what activities, services and products are to be provided in return for a benefit, usually money. Performance measurement as defined in this paper, is used in OBA, monitoring, and contracting, specifically in its use to document outputs, outcomes, or compliance.

Throughout the literature, the reason stated for using performance measurement was because a funder required it (Carmen & Fredricks, 2010; Carnochan et.al, 2014; Christensen & Ebrahim, 2006; Dehoog, 1985; Hoefler, 2000; MacIndoe & Barman, 2012). These studies included an array of different types of human service organizations including mental health, child welfare (Carnochan et.al, 2014), and those that provide basic needs such as food, clothing and shelter (Christensen & Ebrahim, 2006; DeHoog, 1985).

Dicke (2002) found that 98% of those surveyed said monitoring was the most common method of ensuring accountability to funders and auditing was second at 88%. Outcome based assessment, or performance measurement, was third at 85%. What is interesting in this study is that 100% of the funders in the study claimed that monitoring was the most effective way to ensure quality services were being provided, with outcomes being second at 58%. Only 2% of the case managers trusted monitoring and 58% of case managers questioned the validity of information being provided to the funders (Dicke, 2002). Dicke (2002) surveyed and interviewed staff of both funding organizations and provider organizations. Staff members at the provider organizations included management and caseworkers.

Using performance measurement for control and conformity is not without challenges (Carnochan et al. 2014). Using the terms “performance measurement” and “performance management” interchangeably, implying the context was as control and monitoring (Carnochan et al. 2014). Their study found four challenges related to defining outcomes: a) complexity of client goals, b) how to handle aggregation of data, c) outcomes were not systematically defined, and d) disconnect between funder mandates and staff assessment of client progress. Nonprofit managers needed quantifiable data, but caseworkers documented the progress of clients through case notes that were not captured in agency level reports. Carnochan et al. (2014) found that there seemed to be a consensus by both management and staff that funder outcome reporting was leading the push for measurement, and that performance measurement (data collection), was limited to what was required by funders.

Performance Measurement for Improvement

One research question of this study is whether front line human service workers view performance measurement as a tool for program improvement or as a compliance activity for funder or regulatory requirements. Some of the research on human service management tries to determine whether nonprofits use performance measurement as a tool to improve programs or organizational effectiveness (Carnochan et.al, 2014; LeRoux & Wright, 2010; Thomson, 2010). Carnochan et al. (2014) make a point to distinguish between performance measurement and program evaluation, noting that while they are different they can be complementary. Evaluation can provide information to improve outcome measurement and performance measurement can aid in evaluation design. Other studies cite the limitations of performance measurement in improving services, such as lack of baseline data (Dicke, 2002) needed for program evaluation.

The work of Dicke (2002) points to the use of rules and monitoring as external controls to limit the discretion of the subcontractor, and thereby the caseworker (Dicke, 2002). O'Neill (2002) provides an explanation of how compliance monitoring can impede improvement in service delivery:

Perhaps the culture of accountability that we are relentlessly building for ourselves actually damages trust rather than supporting it. Plants don't flourish when we pull them up too often to check how their roots are growing: political, institutional, and professional life may not flourish if we constantly uproot it to demonstrate that everything is transparent and trustworthy (O'Neill, 2002, p. 19).

Thomson (2010) found that despite resource constraints, performance measures increased as reporting mandates increased; they were not used as a tool for improvement. Hoefler (2000) found that the main reason for evaluation was for internal control (57 %), followed by curiosity about their program effectiveness. Only 39% of organization leaders indicate it is funding requirements that drive evaluation (Hoefler, 2000). This appears to contradict Thomson (2010).

Literature from the evaluation field clarifies some of the issues with using performance measurement for improvement. Liket et al. (2014) look at performance measurement from an evaluation perspective. Their findings are that nonprofits have difficulty with sound evaluation when coupled with accountability. Liket et al. (2014) tested a framework that involved collaboration between the funder and the provider. The stakeholders in this study were not concerned about control and monitoring as much as publicizing the success of the program. One conclusion Liket et al. (2014) reached about the study was that nonprofits were not able to use

evaluation successfully to meet their own organizational goals if they were not part of developing the evaluation process.

Organizational Learning and Feedback

An aspect of using performance measurement for improvement is understanding how improvement happens within an organization. *Organizational learning* is a term used to describe “a change in the organization that occurs as the organization acquires experience” (Argote & Miron-Spektor, 2011, p. 1124). These changes include adapting to additional work load that is created by performance measurement and changing organizational processes to adjust to performance measurement requirements (Benjamin, 2008). Organizational change is synonymous with flexibility and organizations need to be flexible in order to respond to external events in unstable environments (Levitt & March, 1988). Whether it is competition for resources, changing technology, or political activism, leaders are asked to make changes in their organization, this calls for changing strategies, changing power structures, relearning and other adjustments to the internal functioning of the organization (Yukl, 1999).

Research in organizational learning has identified varying elements necessary for an organization to be a learning organization, but one element that is consistent throughout the research is feedback (Argote & Miron-Spektor, 2011; Argyris, 1976; McElroy, 2000; Senge, 2006). Feedback generally means the transfer of information between two entities for the purposes of understanding the effectiveness of an action (Argyris, 1977; Ashford & Cummings, 1983; Senge, 2006). Feedback allows an individual or entity to adapt behavior toward improvement or to reinforce current practices that are working (Kazbour, McGee, Mooney, Masica, & Brinkerhoff, 2013). Simply sharing information to another persons or entity does not constitute feedback, an analysis of that information needs to be returned to allow the information

to be used to reinforce or change practices and behaviors. Argyris refers to this process as double loop learning (Argyris, 1977). Research on feedback has shown that it is effective in increasing organizational performance (Kazbour et al., 2013) as well as employee performance (Jacobs & Washington, 2003). Considering the known effects of feedback coupled with the use of feedback for organizational learning, and the practice of performance measurement for accountability; the quality of the feedback that occurs through performance measurement can assist in understanding the effects of performance measurement on organizational effectiveness.

Human Services as Gendered Labor

Human services organizations are “symbols of the caring society a manifestation of the societal obligation to the welfare and well being of its citizens” (Hasenfield 2010, p.10). Human services as a field is broad with the primary function of human services to “Assist individual and communities to function as effectively as possible in major domains of living” (National Organization for Human Services, 2018). These services are so diverse they include childcare to assisted living facilities and correctional institutions to food banks. The common theme throughout human services is that they assist people in meeting the needs and challenges of their everyday life. Some goals require a change in behavior of the client, such as reduced drug use. These changes are often difficult and uniquely challenging to each client, and often requires years of work.

Because of the caring nature of the human services field, it often requires *emotion management*. In 1983, Hochschild contended that gender is a factor in women’s need to commoditize emotion work since they had less access to other resources such as money and power (Appelrouth & Edles, 2011). Emotional labor is a concept developed by Hochschild (Appelrouth & Edles, 2011; Grandey, 2000; Kruml & Geddes, 2000) to describe the individuals’

selling of their ability to act in emotionally appropriate ways as constructed by employer expectations. In this way employees help to meet organizational goals (Grandey, 2000). Emotion work is required; the “act of trying to change in degree or quality an emotion or feeling” (Hochschild [1983] in Appelrouth & Edles 2011, p. 245)”. This is how individuals commoditize their ability to manage emotions to employers for income.

There are three characteristics of labor acts that required emotional labor; a) direct contact with the public through face-to face interaction or directly talking to a person, b) the requirement by employers to lead clients to a particular emotional response, c) control of employees’ emotional manners through training and supervision (Appelrouth & Edles, 2011; Grandey 2000). “Success within the ‘people professions’ requires diligently controlling one’s emotions in order to express the appropriate feelings state (trustworthy, dedicated, ambitious, caring etc.) associated with the human services occupations” (Appelrouth & Edles, 2011, p.238).

Another concept important to understanding gendered labor is feeling rules, also developed by Hochschild (Appelrouth & Edles, 2011). Feeling rules are generally acceptable emotional responses to a situation. They are socially constructed and our own epistemological expectation of inner emotions determines how much emotional work is required to maintain emotionally appropriate responses and activities. Rules include extent, duration, and direction of feelings (Hochschild [2003] in Appelrouth & Edles, 2011). Society has feeling rules to tell us when it is appropriate to cry, how hard, and how long before it becomes socially awkward. Feeling rules tell us how to appropriately manage anger in a social setting. Our own epistemological belief in our ability to manage feeling rules may determine what movies we want to see to avoid crying in public, or attendance at a party to avoid specific interpersonal interactions. Feeling rules “tend to be latent and resistant to formal codification” (Hochschild

[2003] in Appelrouth & Edles, 2011, p. 247). They also reflect “patterns of social membership” with some being universal and some being endemic of a specific social group (Hochschild [2003] in Appelrouth & Edles, 2011, p. 247). Hochschild also points out that feeling rules, and our individuals choose to follow those rules, are tied to our ideological beliefs. Thirty-five years after Hochschild penned *The Managed Heart*, she notes that things have changed: American feminism has given women the legitimate right to respond with anger at work abuses and men can care more about work-life balance (Hochschild [2003] in Applerouth 2011,). Yet the work human service is still largely a female occupation and emotional work can still take its toll.

There are two schools of thought on how emotional labor affects the employee (Grandey, 2000). Hochschild contends that emotional labor can result in burnout or stress. But the work of Ashforth and Humphreys (1995) claim it could be effortless (Grandey, 2000). The difference is in whether the employee is managing emotions in a manner that is genuine or if those emotions cause a disassociation with the true feelings of the employee. Grandey (2000) found that research supports that there are health risks associated with emotional management that include cancer, high blood pressure, and heart disease. Grandey (2000) also found that research suggests that support from coworkers and supervisors can mitigate the stress caused by emotion management. This makes perspective of receiving support and feedback important to overall employee satisfaction and performance.

Although it is not intent of this study to provide in-depth analysis of the emotion management or gendered labor as a factor in organizational effectiveness or performance measurement; it is a valid concern since this study takes place in the human services field. There may be some evidence of biases or practices that are influenced by gendered issues.

Emerging Areas of Research

Christensen and Ebrahim (2006) examined how accountability affected the daily operations of one of 450 offices of a national nonprofit that provides services to 30 million immigrants and refugees. This is a single focus agency with a religious affiliation that receives primarily government funding. This study assumed that there were competing demands for accountability and it focused its investigation on how staff negotiated the tension created by these demands. One example of the onerous reporting was the process of literally counting spoons in the residence of a refugee to document all of the resources provided to the client. Much of the success of negotiating the tension created by accountability demands is attributed to the director. She “has set up buffers to protect staff, to some degree, from these demands” (Christensen & Ebrahim, 2006, p 201). She also negotiates with funders to reduce the redundancy in reports. Using an open-door policy, communication strategies, and clear focus on client needs, the director has created a climate that seems to buffer the negative effects of arduous performance measurement requirements. This suggests that leadership may be an important variable in understanding management and staff perceptions and attitude toward accountability practices. This idea is also supported by Packard (2009) who finds that leadership is one of the top ten most important factors for achieving desired outcomes. The importance of leadership in mitigating possible negative affects of accountability measures is supported but the work of Brown et al. (2015) who found that NPO executives valued human capital more than performance measurement in increasing organizational effectiveness. This extends to the caseworker as well.

Another emerging area of research is on the interaction between the client and the direct service staff (Benjamin, 2008, 2012; Benjamin & Campbell, 2015). The focus of this research is the relationship between funding requests for outcome measures and the accounts provided by

caseworkers to describe the work they do. Benjamin (2012) did a qualitative study that looked at the existing outcome measurement guides on performance measurement and compared the required measures to how front-line staff describes their work. Benjamin (2012) reviewed studies that included first-hand descriptions of what front line staff members do as part of their work in human services, neighborhood development, as well as grass roots organizations. Benjamin (2012) contends that a disconnect exists between what frontline staff members identify as making their work effective and what outcomes measures funders are requiring them to use. These case studies actually involved interviews with various levels of staff, including volunteers and executive directors (Benjamin, 2008). There is no clear detail in Benjamin's 2012 study as to which interviews were used nor is there a distinction between the viewpoints of the managers from front line service delivery persons in his analysis.

An extension of this work is the research done by Benjamin and Campbell (2015). They conducted interviews with frontline staff in eight human service nonprofits. Their findings are that front line staff does *co-determination* work to affect change in clients' lives. The goal of this study was to inject the role of clients' agency in the performance measurement framework. Findings suggest that there are four common relational practices: listening, naming, challenging and linking. The conclusion is that the performance measurement focuses on activities completed, or outputs, and not the changes for the client.

Benjamin and Campbell's (2015) purpose is not to dispute the importance of using performance measurement, but to demonstrate that they typically do not go far enough because performance measurement often omits important client outcomes. Program interventions appear to be key in producing desired client outcomes, overshadowing the importance of the inter-relational aspects between the client and the worker (Benjamin & Campbell, 2015). Packard's

(2009) research supports this. He found that the satisfaction of both the client and the employee ranked higher than client outcomes in importance for organizational effectiveness.

Limitations of Existing Literature

Several weaknesses in the body of research concerning performance measures in nonprofits have been identified. The nature of nonprofit organizations themselves adds to the complexity of the topic. Some researchers have narrowed down the types of nonprofits by omitting arts and museums (MacIndoe & Barman, 2012) but little distinction is made in the literature between human service nonprofits and nonprofits in general. Even when narrowing down the type of nonprofit to just human services, there are large differences within the types of human service agencies (Herman & Renz, 1997; LeRoux & Wright, 2010; Thompson, 2010). A human service nonprofit can be a hospital, childcare center, food pantry, or domestic violence shelter. Many nonprofits provide a wide range of services within their own organization. This has implications for comparing staff and management perceptions when education, professionalism, and work requirements are factors in understanding attitude and perceptions.

Another difference in some nonprofits is there are some with a single focus and others with a multi-focused mission. For example, Child and Youth Services agencies typically focus on one goal, the health safety of the child. Child and Youth Services are government human service agencies, and some nonprofits are contracted to provide these services. But not all human service agencies are single focus agencies. Community Action Agencies (CAA) are one example. A CAA is tasked with addressing poverty and may work on homelessness, employment, head start, weatherization, food insecurity, adult literacy, and other identified community needs. This variety in NPOs suggests a need for focused research that takes into account the variations in the

nonprofit field that extend beyond structure, size, and age (Glisson & Martin, 1980; MacIndoe & Barman, 2013).

The lack of differentiation among aspects of nonprofits adds to other challenges in understanding the impact of performance measurement on NPOs. Researchers approaching the topic from an organizational level often did not articulate the differences in the organizational dynamics of the funder versus the provider. Some research placed the organization in the role of direct service provision rather than the staff member that provided the actual service (Carachan et al., 2014; Dicke, 2002). By not making distinctions between the caseworkers and the executive director in these studies, the voices of both the management and the worker were presented as one.

Another lack of distinction was the work of the nonprofit itself. Much of the research approaches all human services nonprofits as the same, irrespective of the mission (Carachan et al., 2014; MacIndoe & Barman, 2013). Studies looked at nonprofits that provide mental health, child welfare, disability services, and homeless prevention as if the characteristics of the workers would be the same when the qualifications of workers range from licensed professionals to high school graduates (Carachan et al., 2014; Dicke, 2002; Thomson, 2012).

Theoretical Framework

There are several theoretical frameworks which can guide research on the effect of performance measures in nonprofit human services. The metatheoretical perspective of structural functionalism encapsulates many of the questions I explore in this study. Concepts appropriate to structural functionalism include “functional need, dysfunction, feedback, institutionalization” (Powers, 2010, p.152) among others. Structural functionalism can be described as a research lens through which “social *structures* can be understood by studying the *functions* they perform for the

maintenance of the subsystems of which they are a part” (Tompkins, 2005, p. 213). In this study the social structures are the nonprofit organizations, and performance measurement is a function that is performed in those organizations.

Weber’s bureaucracy is integral to understanding how these mechanisms work together in the context of organizational structures as they relate to practices and problems of accountability within large institutions such as the federal government. This discussion will include a basic summary of bureaucracy and how Merton’s ideas about latent and manifest functions as well as strain theory help inform the processes of performance measurement. Merton’s idea of latent functions can be used to understand the unintended consequence of performance measures on human services delivery. Strain theory can be used to understand the staff member reactions to the bureaucratic requirements involved in performance measurements. Institutional theory (Jaffe, 2001) can be used to understand the process by which the use of performance measurement has become prolific by funders and organizations. Institutionalization will be discussed in relation to how performance measures have become a standard practice even though their validity and usefulness in the decision-making process are questioned (Moynihan, 2006a).

Bureaucracy

A bureaucracy is a highly organized structure that is dependent on rules and procedures to remain objective and fair as well as provide for efficiency (Tompkins, 2005). Weber (Weber, 1914/2007) developed the concept of ideal-type of bureaucracy in his efforts to objectively describe and analyze administrative systems. Ideal-types do not signify the best type, but they reflect a standard that can be used for comparisons and help to communicate sociological thought and research. Bureaucracies are a means for controlling and organizing legitimate power (Tompkins, 2005). The characteristics of an ideal-type bureaucracy are clearly defined rules and

duties, hierarchical authority structure, systems of rules that limit personal discretion, promotion and reward based on technical merit over progeny or favor, and documentation (Tompkins, 2005; Udy, 1959). Although each of these characteristics holds a purpose in large organizations, the limit of personal discretion is one aspect that can dampen the ability to provide for the unique needs of human services clients.

For Weber, bureaucratization represented a legal-rational solution to the power relationship of authority to citizens (Tompkins, 2005). A bureaucracy's power is dependent on the perception of legitimacy. This ideal-type of organizational structure lends itself to reliable behavior because decisions are made according to preplanned procedures that limit individual decision-making (Kuhn, 1971). This reliable and predictable behavior is appealing if the goal is to operationalize activities for evaluation purposes. Performance measurement is an example of a bureaucratic response to the call to reduce the size of government by providing a procedure that limits bias in the evaluation of a program's effectiveness. In human service organizations, adherence to rules extends not only to simple tasks such as determining eligibility and filling out forms but also includes program implementation. On one hand, this promotes impartiality and fairness for the clients. At the same time, it challenges the ability of workers to make decisions based on the individual and unique needs of the client while being compliant with funding requirements.

Gouldner expanded the concept of bureaucracies into ideal-type patterns of management (Tompkins, 2005). One ideal-type management structure is where failure to obey rules leads to punishment or a punishment-centered bureaucracy (Jaffee, 2001). Another ideal-type allows for the participation of workers in the development of procedures and processes which is called a representative bureaucracy (Jaffee, 2001). In bureaucracies, workers are encouraged to accept,

through training and recognition, the greater knowledge of authority; the expectation is for staff to adhere voluntarily to the rules (Jaffee, 2001; Tomkins 2005). This study explores the type of communication that occurs between the front-line staff and management and the perception of how that communication creates a sense of risk or an environment of partnership in the process of performance measurement.

Bureaucracy is not currently viewed as the best way to organize organizations for effectiveness (Jaffe, 2001). Merton (1940) identified some of the dysfunctions of bureaucracies. His view was that the results of bureaucracy were “institutional rigidity, goal displacement and disgruntled clients” (Tompkins, 2005, p. 215). Merton provides a framework for understanding the unintended consequences of bureaucratic actions through his concept of manifest and latent functions.

Functionalism and Manifest and Latent Functions

Robert Merton suggested the use of the terms manifest and latent functions in 1936 as a means for systematic investigation of the latent or unintended consequences of manifest or intentional actions (Merton, 1936). Merton did not lay claim to the origination of the concept but determined a need to clearly define the concept so it could be further investigated and expand social inquiry. Merton’s definition (1957/ 2011) of manifest and latent functions is

The first referring to those objective consequences for a specified unit (person, subgroup, social or cultural system) which contribute to its adjustment or adaptation and were so intended: the second referring to unintended and unrecognized consequences of the same order (Merton, 1957/ 2007, p.441).

One use Merton states for making this distinction is to direct the attention of the researcher away from intended consequences to unintended consequences. If a researcher only studies the intended

consequence (manifest) then the researcher risks only supporting the status quo (Merton, 1957/2007). If a researcher focuses on unintended consequences, then s/he is open to a larger area of inquiry to answer questions such as opportunity costs and other effects of an action that are not directly associated with the activity or the intended consequences. This concept is especially pertinent when considering Merton's views on bureaucracy. Merton stated in 1957 that "bureaucratization entails the separation of individuals from instruments of production" (Merton, 1957/2011, p. 108). When funders requiring performance, measurement are separated from the caseworkers, and decisions are made by looking at numbers and not the many variables in actual events then the separation described by Merton occurs. The unintended consequences of the performance measurements on service delivery should be considered when looking at their effects on program delivery.

When considering bureaucracies, Merton notes that "if the bureaucracy is to operate successfully, it must attain a high degree of reliability of behavior, an unusual degree of conformity with prescribed patterns of action," (1957/ 2011, p.109). This can apply to performance measures: if information is to be aggregated from multiple agencies and organizations, then it has to be conforming to prescribed patterns and activities in order to be reliable. This begins the process described by Merton where goal displacement occurs; the adherence to the rules or procedures is more important than the actual goal of the activity. Merton writes that goal displacement leads to adaptation when the present circumstance does not align with the existing policies and then making "the very elements which conduce towards efficiency in general, produce inefficiency in specific instances" (Merton 1957/ 2011, p. 110).

Strain Theory

Merton provides the framework for strain theory in his 1938 article *Social Structure and Anomie* (Agnew, 1992; Farnworth & Leiber, 1989). His discussion describes two components of social structure, “cultural goals and institutional norms” (Merton, 1938). The inability to reach goals through legitimate means leads to adaptation. In Merton’s discussion individuals adapt to a conflict between goal and acceptable practices include conformity, innovation, ritualism, retreatism, and rebellion (Merton, 1938). These adaptation responses constitute a continuum the individual may take ranging from completely ignoring institutional norms to obtain the goal to completely abandoning the goal to adhere to institutional norms. This is not a static state in the individuals may vary according to circumstances and means (Merton, 1938). Merton describes the complete abandonment of the institutional structures or rebellion:

When both the culture goals and institutionalized procedures have been assimilated thoroughly by the individual and imbued with affect and high positive value, but where those institutional procedures which promise a measure of success attainment of the goals are not available to the individual. In such instances, there results a twofold mental conflict insofar as the moral obligation for adopting institutional means conflicts with the pressure to resort to illegitimate means (which may attain the goal) and inasmuch as the individual is shut off from means which are both legitimate *and* effective. (Merton, 1938, p 677).

Contemporary researchers have developed a General Strain Theory postulates that there are stressors can lead to behaviors that are not generally acceptable (Agnew, 1992; Broidy & Agnew, 1997; Farnworth& Leiber, 1989). The catalyst for these behaviors is emotion (Agnew

2001). Agnew defines *subjective strain* as “events or conditions that are disliked by the people who are experiencing (or have experienced) them” (Agnew, 2001, p.321). Subjective strain is subjective because, as Andrews (2001) contends, individuals react differently to subjective strain with different emotions and intensity of emotions. There is also some research that suggest that there are differences in response to subjective strain associated with gender (Broidy & Agnew, 1997).

A factor to consider in an individuals response to subjective strain is social control (Agnew, 2001). Working conditions associated with struggling to meet goals or compete tasks are often associated with high social control (Agnew, 2001). This high social control means that individuals do not resort to crime, but to alternative strategies to alleviate strain (Agnew, 2001). Goal displacement, strain theory and emotion management are all concepts to consider when human service workers are faced with providing services to clients within the context of rigid program requirements. General strain theory suggests that caseworkers will seek alternative solutions to helping the client. But their willingness to work outside the rules may depend on their decision to abandon the goal of human services to help the client versus accepting the institutional norms that are requiring adherence to performance measurement requirements. Research in emotion management suggests that a high degree of social control felt by the caseworker may also affect the decision to work outside the rules.

Bureaucracy as a Gendered Organization

Organizational theory has held on to the notion that organizations are gender neutral, (Acker, 1990). However, Ackers, a social feminist theorist, (Acker, 1990) proposed that organizations are gendered. A gendered organization is described as a masculine or oppressive organizational structure, such as bureaucracy. This structure is due to male power domination and

rational rules and procedures (Acker, 1990; Britton, 2000). Abstract conceptualizations of seemingly gender-neutral job descriptions perpetuate the notion of a gender-neutral organization when such conceptualizations require a masculine availability of the job and deny a women's essential feminine reproduction and emotional qualities. The result is the feminization of the people within bureaucratic organizations because they become stereotypically passive and powerless (Acker, 1990). What Aker explores is how

Masculinity always seems to symbolize self respect for men at the bottom and power for men at the top while confirming their gender's superiority.

Theories that posit organizational and bureaucracy as gender neutral cannot adequately account for this gendered structuring (Acker, 1990, p.145).

Other theorists have suggested that gender is not limited to a masculine interpretation of gendered organizations but should reflect either a masculine or a feminine organizational structure (Britton, 2000).

To understand an organization or occupation as gendered is to say that an organization or occupation has qualities and structures that make it identifiable as either masculine or feminine. That conception of the occupation, as either feminine or masculine, attaches value to that occupation or organization and the differences in resources and status that are aligned with gender (Aker, 1990; Britton, 2000). The concept of gendered also suggests an iterative process whereby the gender we display in our jobs affects that job and the gender expressed by the job affects our personal attributes (Britton 2000). At the most basic level, a gendered organization is such by virtue of the predominate gender that works in that occupation (Britton, 2000). This can be an oversimplification that gender typing addresses. Gendered occupations are occupations perceived as more appropriate for individuals with masculine or feminine characteristics (Britton, 2000).

Institutional Theory

Two foundational works in Institutional Theory are Meyer and Rowan's 1977 article *Institutional Organizations Formal Structure as Myth and Ceremony* and DiMaggio and Powell's 1983 article *The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields*. Institutional theory, also referred to as institutional isomorphism, looks at the processes by which organizations become homogenized through the life cycle of an organizational field (DiMaggio & Powell, 1983; Suddaby, 2010). Institutional theory acknowledges the environmental forces that play on the structures and practices of organizations (Jaffe, 2001). "Something outside the boundaries of the organization influences internal operations" (Jaffe, 2001, p.209). When organizations become similar, they are adopting practices and structures that are expected within their fields. According to Meyer and Rowan (1977) the reasons to adopt accepted policies and structures are to gain legitimacy, increase stability, and obtain resources to survive. Institutional theory can inform the adoption of performance measures by NPO's from several perspectives: the pervasive use of outcome measurement in areas where they are not the best way to measure effectiveness and conformity and the practice of outcome measures by authoritative entities that do not feel the effects of their requirements (Pfeffer & Salanick, 1974).

Research in institutional theory has explored why organizations adopt practices that do not necessarily make sense from a purely economic perspective. These practices create an air of legitimacy but aren't congruent to organizational goals. Research also describes how processes and outputs are instilled with meaning that exceeds their value (Suddaby, 2010). DiMaggio and Powell (1983) suggest three processes that drive institutional isomorphism: coercive, mimetic, and normative processes. The coercive process is when forces of authority outside the

organization drive the adoption of practices, often through legislation or policy requirements. Mimetic isomorphism is the process by which organizations pattern their structures and practices on other organizations that they deem successful. Normative processes are explained through professionalism: established credentials or coordinated behaviors within a profession that cross organizational fields.

Powell and Colyvas (2008) posit that a micro-level theory of institutionalization would help inform the macro level analysis by looking at how individual actions affect the structures of organizations. For example, one response to coercive institutionalism is decoupling activity from formal structure (Meyer & Rowan, 1977). This occurs when organizations have a set of practices that meet external demands but also have an informal structure that allows activities to be effective. This process could help inform the actions of caseworkers when policies and procedures do not reflect the actual lives of their clients.

Chapter Summary

The climate in both government and foundations has been heavily focused on proving that the resources they spend on human services are used effectively: funders, and society at large, demand accountability. One aspect of accountability for organizations is showing effectiveness in meeting their goals and one tool for showing that effectiveness is through performance measurements. At the same time that the need for accountability has increased; nonprofits have played an increasingly large role in helping communities meet the needs of low income individuals and families through government programs.

Demonstrating effectiveness is difficult when dealing with complex problems such as poverty. Performance measurement attempts to show what a program has accomplished through a collection of a variety of data and reporting outputs and outcomes. Some researchers question the

value and validity of information gained from performance measurement (Dicke & Ott, 2002; Ebrahim & Rangan, 2010; Pandey et al. 2006). Stakeholders also need to understand how performance measurement can be translated into organizational effectiveness: is this practice actually leading to improvement of programs and organizations.

As organizations institute performance measurement to show they are effective and therefore accountable, these policies can affect the job of the caseworker. This study uses qualitative approach to understand the perceptions of caseworkers at NPOs. Managers and executive directors are also included to provide some additional insight into the impact of performance measurement in their work and organizational performance. The research questions will focus on three areas of concern:

Research Question 1: How do nonprofit program staff members perceive performance measurement and its impact on their job?

Research Question 2: Do staff members feel they have adequate training and time to provide quality data.

Research Question 3: What type of communications exists concerning performance measurement results for the improvement of programs?

The caseworker is the cornerstone of an organization's ability to deliver services to clients effectively (Alexander & Charles, 2009). Performance measurement in nonprofits, which literature suggests is primarily used to satisfy accountability requirements of funders (Hoefler, 2000; Thomson, 2010), affects the work of human service providers. There are complex relationships between the organization, the funder, the client, and the caseworker. The caseworker can react to goal displacement, strain, or the need for emotion management in a manner that reflects on the ability of the organization to be effective within any aspect of their work in human

services. This study provides a focused look at the impact on the caseworker of performance measurement practices within the larger context of their work in the human services field.

CHAPTER THREE

METHODS

This study used qualitative methods to explore the perspective of staff members in human service nonprofit organizations (NPO's) on performance measurements. Included in this chapter is the research design that describes the methods used and rationales for them; a description of how the research sample was selected, both the site selection and the participant selection; data collection procedures and ethical considerations. I also provide a description of the data analysis process I used, issues of trustworthiness of the data and results, as well as the limitations of my methods.

Research Design

There are two broad strategies that can be used to approach a research problem: qualitative or quantitative research. When researchers find there is not enough information on a topic to develop consistent operational definitions of the concepts being studied or the precise hypothesis (Monette, Sullivan, & DeJong, 2011), it is appropriate to use qualitative research strategies. That is the case in this study. Qualitative research uses data in the form of words, narratives, and descriptions (Monette, et al., 2011). Quantitative research tests theories using variables that researchers can measure for statistical analysis (Creswell, 2009). One factor I considered when deciding the research approach is the amount of information available the topic. Another was the type of information that I sought. I chose a qualitative design because there is very little information in the literature on the perceptions of caseworkers regarding performance measurement and because I wanted in-depth information on this subject. According to Maxwell (2005), qualitative methods are particularly suited to understanding the "meaning" an activity or event has for the participant. Qualitative research also focuses on the meanings of the

phenomenon to the participant, not the meaning to the researcher (Creswell. 2009). This research topic, staff perceptions of performance measurement in nonprofit organizations, spans the public administration field, nonprofit literature, and evaluation literature with each field using a different perspective. Because of this lack of shared meaning about the topic, there is disjuncture in the literature, and a qualitative research design has allowed me to uncover perceptions of workers without relying on all participants understanding the language associated with performance measurement.

My research design used semi-structured individual interviews of caseworkers, program managers, and agency executives at seven different agencies. There are several strengths to qualitative interviews as opposed to a quantitative method for my purposes. According to Monette et al. (2011), participants in interviews are more motivated to provide responses and there is an opportunity to explain questions. An interviewer has more control over the process than with more structured and quantitative collection methods. The interview also allows for direct observation of the participant during the interview to record and analyze nonverbal expressions. The disadvantage of the interview is the time involved for both the participant and the interviewer (Monette, et al. 2011).

I recruited participants through purposive sampling. Using the Community Action Agencies list for Pennsylvania, I identified employment and training programs as the sample sources for the study for several reasons: staffing these programs does not generally require a specialized degree or certification and these are common programs within the Community Action Agencies. Although employment and training programs have program guidelines, there are not the strict protocols that are required of other types of programs that come under the “evidence-based” category. This allowed me to compare staff perceptions that include different ways of

implementing performance measurement, while at the same time keeping some factors that would influence their perceptions consistent.

I do not intend to identify or assert causal relationships from the data I collect, as might be done in a quantitative research. Instead, my aim is to look in depth for patterns that might emerge and highlight connections (Creswell, 2009), among aspects of performance measurement and dimensions of the work of human services. Drawbacks to qualitative studies like this one are that some concepts and perceptions are not easily quantifiable. However, when a qualitative method is used, it provides more detailed and precise information about human behavior (Monett, 2011). The goal is neither to quantify nor generalization but to explore and describe a phenomenon, particularly one we know little about, such as the focus of this research.

Research Sample

Typical of qualitative research, the sampling was purposive (Bloomberg & Volpe, 2008). The research sample used two layers of criteria. First, I selected Community Action Agencies (CAAs) that offer employment and training programs. Second, I recruited staff members who work as front-line service delivery caseworkers, program managers, and agency directors as participants for interviews. Since this study is mainly concerned with the caseworkers' perspective, I did not intend to have a large number of interviews with management personnel.

Site Selection

This study explores perceptions of staff at nonprofit organizations, but there is a wide variety of nonprofits even within the human services field. Community Action Agencies (CAAs) are a network of federally funded agencies that provide services to low income clients. In Pennsylvania, there are 41 such agencies that cover all 67 counties (<http://www.thecaap.org>). The rationale for this pool of organizations is both the uniqueness of CAAs and the shared

performance measurement processes they use. Each CAA provides a different array of services, but the reporting requirements for the agency are the same across the state and are dictated by federal legislations such as the Government Performance and Results Act of 1993 (United States General Accounting Office, 1997).

The Community Action Association of Pennsylvania (CAAP) website, www.thecaap.org, provided information on all CAAs in Pennsylvania. I obtained a listing of all programs provided by each agency dated from 2009 and updated in 2016. I attempted to contact each agency by phone to update information on the Executive Director and to confirm that the agency did implement employment and training programs. I found 34 CCA's that met the criteria and I invited 19 agencies to participate in the study based on willingness to respond to initial contact. Using this list, I invited all executive directors at all 19 CAA's by email or phone to participate in the study. Executive directors at nine agencies gave verbal agreement to participate in the study. Two agencies agreed to participate but did not return written permission, so they were not included in the study.

Participant Selection

Once executive directors agreed to allow their staff members to participate in the study, they provided a contact person at their agency to coordinate with the agency staff for interviews. Since employment and training staff members were small in number at each agency, I requested the agency contact person to allow me to contact their employment and training staff directly through phone or email. I coordinated a date at each agency for face-to-face interviews with the contact person, and, upon arrival, staff members volunteered for interviews that were available on that date. Phone interviews were available as an option for staff members that were not available

to interview face to face. I provided all study participants with the IRB Informed Consent form prior to interviews through email.

Data Collection

I conducted 28 interviews of staff members who provide employment and training (E&T) programs in seven Community Action Agencies. Study participants signed an Informed Consent form, and I also reviewed it with them. I conducted 24 interviews in-person and four interviews by phone. All study participants gave permission to record the interviews.

In-person interviews took place during work hours in a private office location at the agency. I took written notes during the interviews to record nonverbal cues and details of the setting. The average interview length was 44 minutes, and the median length of the interviews was 38 minutes. The longest interview was 95 minutes, and the shortest interview was 22 minutes (excluding an interview that the study participant halted at 10 minutes).

The interview questions I developed were intended to facilitate a conversation interview and to prompt study participants to share their experiences that pertain to the research topic. I conducted the interviews in a conversational style and this meant that I sometimes asked some interview questions out of order, as can happen in the flow of semi-structured interviews. Study participants also provided responses that touched on multiple topics within one response.

Although interviews were conversational in nature, I used a semi-structured interview guide sheet when interviewing to ensure consistency across interviews to avoid variation in the experience from one participant to another. The semi-structured interview guide includes the order of question as follows:

1. As part of your job, I understand that you collect information about your employment and training clients and programs. What's your experience with that?

- a. Tell me about the types of information and how you collect it?
 - b. Do you run into any difficulties gathering the necessary data?
2. How does collecting this information affect how you do your job?
 - a. Tell me about the time it takes?
 - b. How does that impact the help you can give clients?
3. What's your sense of how this information gets used in terms of programs for service delivery to clients?
 - a. What types of reports are generated from the information?
 - b. How is funding affected?
4. Tell me about the training you get for collecting and using this information?
 - a. What types of trainings do you wish you received?
5. Tell me about how you are involved in the process; what input do you have?
 - a. Do you track more information than is required?
6. From your perspective, what would it take to make the information you collect more meaningful?
7. What type of feedback do you get from the information you collect?
 - a. What type of feedback would you like to get?

The semi-structured interview guide is included as Appendix A. It contains the heading, date, time, standardized introduction, list of questions and probes with space to note nonverbal participant actions such as pauses, shrugs, facial expressions, or laughter. I asked the demographic questions at the end of each interview.

All interviews were transcribed by me verbatim. In order to learn from experience, I began the transcription process while still in the data collection phase of the study, before all interviews were complete.

Ethical Considerations

As Creswell (2009) notes, in qualitative research, the researcher's "interpretations cannot be separated from the own backgrounds, history, contexts, and prior understandings" (p. 176). Traditionally, experience can be perceived as a bias in research that needs to be avoided, but Maxwell (2005) views a personal background in a study as an asset that informs insights, can assist in the development of hypothesis, and helps with validity checks. As an employee of a Community Action Agency that has direct responsibility for reporting and data collection, I have worked under the same requirements experienced by the study participants, but I also understand that the ways individuals and organizations handle these requirements will be different.

My experience also includes extensive training in CAA performance requirements. I am a Nationally Certified ROMA Trainer (NCRT), which qualifies me to train CCA staff, administration, and boards of directors on the Result Oriented Management and Accountability (ROMA) process used by CAAs. The in-depth knowledge I have of the performance measurement philosophy and processes used by CAA's, coupled with firsthand knowledge with the implementation of the practice at the organizational level, led to my interest in this topic. I did not share my qualifications with the study participants so as not to bias their responses. I understand that the practices of CAAs have implications on a variety of levels. I also was aware that I had limited knowledge of how these performance measurement practices impact direct service delivery or whether the agencies or funders use the data to improve service delivery. Consequently, by studying several similar agencies with similar programs, my goal was to gain

insight on how service provision among federally funded programs in general may be impacted by accountability. In no way is this study for advocacy of a particular agenda nor is it a sponsored study. However, I acknowledge that my background predisposed me to expectations, as did existing literature, about what I might find. I guarded against these potential biases by taking as neutral, systematic, and open an approach as possible in collecting and analyzing data.

I developed and followed a protocol for the ethical treatment of participants in my study that was reviewed and approved by the Indiana University of Pennsylvania Institutional Review Board for the Protection of Human Subjects (IRB). The study is low risk, but since I was meeting with participants who would be known to me, and the interviews pertained to their work, the key issue was confidentiality. I received permission from sites through the Executive Directors before I contacted individuals at each site to invite them to participate in an interview. The interviews were voluntary, and the participants signed an informed consent form, included as Appendix B. The transcriptions of the interviews are redacted to eliminate personally identifying information. Participants are unnamed or given pseudonyms when their words are used in reports of this research. No participants withdrew from the study, although one participant struggled to answer questions due to being in a new position, and ended the interview after 10 minutes. IRB Human Subjects Review Protocol Approval Letter has been included as Appendix C.

Data Analysis and Synthesis

According to Welsh (2002), there are three ways to approach qualitative data analysis: literal, interpretive, and reflexive analysis. Literal analysis uses the exact language of participants. Interpretive analysis allows the researcher to assess the meanings implied by the participants' accounts in their analysis. Reflexive analysis is a researcher-centered approach to analysis. In any of these approaches, the data must be organized by coding the text, or breaking it down into

chunks of usable information. For the purposes of this study, I used an interpretive approach because the participants used different language about performance measurement and, although their meanings were typically similar in what they were saying, the wording they used was not consistent. An interpretive approach required me to stay close to the data, identify the meanings participants conveyed, and label those in a way that allowed me to re-use meaning labels across participants when they conveyed the same meanings but used different language to express those means than did other participants.

The coding process occurred in three steps. During the first step, I listened to all the interviews while reviewing the transcripts to get a broad understanding of the responses and the themes that were present. I coded by chunking text and responses into categories that relate to the research questions. Several broad themes emerged in the first step that guided the coding process. These preliminary themes were: 1) the disconnect of the performance measures to program activities, 2) the difficulties that occur when performance measurement is used for many purposes and paired with other processes, and 3) the external influences on performance measurement. Because I conducted the interviews in a conversational style, I asked some interview questions out of the order presented in the semi-structured interview guide, as sometimes happens in the flow of semi-structured interviews. Study participants also provided responses that touched on multiple topics within one response. This resulted in my reviewing each transcript multiple times as patterns emerged so I captured comments made outside the response of specific questions.

For the second step of the coding process, I clustered the interview guide questions that were related to the themes revealed in the initial review of the data. Similar to a code and retrieve system in electronic form, I isolated the responses in a Microsoft Excel spreadsheet to allow me to deconstruct the data in a way that provided insight into how the responses related to a study

participants' position within the agency. This code and retrieve system also allowed me to capture any relevant responses that were not part of a specific interview question. This provided me with an analytic structure that I used to identify the direction and frequency of related responses in each category "across and within groups of questions" (Bradley, Curry & Devers, 2007, p. 120) as well as across and within groups of study participants.

After the initial review and coding of the data, I determined that many of the caseworkers went into great detail describing the actual programs they administered and the process they used for data collection. Although this information was interesting, it did not shed light on the primary research questions about their perceptions of how performance measurement affects the work and the organization's delivery of services. It did reveal, though, that study participants understood performance measurement primarily in terms of data collection and data entry.

The third step of data analysis is finding deeper meanings and connections in study participant responses. As stated by Bradley et al. (2007), "as is true in all qualitative work, our goal was to raise the respondents' comments, whilst respecting the integrity of those comments, to a higher order of abstraction or generality that reflected the patterning seen in the comments," (p. 126). I reviewed responses again to uncover themes and patterns related to staff perceptions of the impact of performance measurement and to expand the initial high-level findings in more detail and specificity. Appendix D provides additional detail of the coding and analysis process.

Issues of Trustworthiness and Limitations

Instead of validity and reliability, qualitative research depends on the trustworthiness of the researcher, which I strive to demonstrate through transparency in my data analysis process (Welsh, 2002). To establish trustworthiness, I have thoroughly documented and described the

coding process. Transcriptions of the interviews as well as any notes or documentation created during analysis are available to the dissertation committee.

The scope of the study is limited in several ways. The sample pool was limited to Community Action Agencies (CAAs) in Pennsylvania to provide a similar organizational background and reduce the amount of organizational variability. Staff members from employment and training programs were selected from within CAA's because those programs are common within CAA's. Sampling employees who work in similar programs in similar agencies reduced the amount of variability within the work of the participants in terms of how they engage with clients and administration. This was intended to make variations in these employees' perceptions about the impacts of performance measurement more clearly attributable to performance measurement rather than variations in program or agency types in which the participants do their work. Although the sample was limited to one program type and one agency type, this strategy provides strength to the study, as well.

Qualitative research provides in-depth understanding of a given topic, but care must be taken in applying the findings in a broader context. Although the results of this study may lend insight to policies and practices, they should not be generalized to other organizations or programs.

Another aspect of the study that narrows the scope is the intent to exclude the general area of evidence-based or research-based programs. These programs are highly structured and field-tested to show program effectiveness. If these programs were included, there would be issues of fidelity as well as a lack of flexibility of the staff in decision-making. While research on the perceptions of staff on the efficacy of research-based programs would be an interesting and worthwhile project, it is not the purpose of this study.

Chapter Summary

Lather (1986) states that “data must be allowed to generate propositions in a dialectic manner that permits use of *a priori* theoretical frameworks, but which keeps a particular framework from becoming the container into which the data must be poured” (as cited in Creswell, 2009, p.65). This study used a qualitative approach to understand the practices and perceptions of performance measurement in nonprofit human services. I interviewed 28 staff members involved with employment and training programs at seven Community Action Agencies across Pennsylvania. Then I organized the data using a multi-stage process to allow for the semi-structured interviews to convey the perceptions of the study participants. Analysis of these perceptions and the descriptions of the practices in meeting performance measurement requirements were organized to address the research questions. This resulted in the findings presented in Chapter Four.

CHAPTER FOUR

RESULTS

The purpose of this research is to understand the perceptions among staff members in non-profit human service organizations about performance measurement. Specifically, I explored the perceptions of performance measurement data collection on the caseworkers' ability to provide services and how they view performance measurement data as it is used in providing services to clients. The research questions focused on three areas:

1. How are performance measurements perceived regarding their impact on service delivery and organizational performance relative to the mission?
2. Do nonprofit organization (NPO) staff members feel they have adequate (a) training and (b) time to collect quality data?
3. How are performance measurement results communicated to caseworkers for improving the delivery of programs and services?

Nonprofit organizations (NPOs) typically implement performance measurements to provide evidence of their effectiveness in achieving the outcomes and goals intended with the funding they receive. Some researchers argue that organizations need to be flexible and responsive to local and immediate situations if they are to be effective (Argyris, 1976; Senge, 2006). In order to make NPOs accountable, some type of performance measurement is necessary (Russ-Eft & Preskill, 2009). There is little research on how the staff members of NPOs perceive the effect of accountability practices on their organizations.

The language concerning performance measurement is used by both program evaluators and by researchers, but does not always have the same meaning (Carnochan et.al, 2014; Hatry, 2002; Russ-Eft & Preskill, 2009). For the purposes of this research, I define performance

measurement as the general activities used to collect and report information about a program's effectiveness and efficiency with the intent of using the information for monitoring, resource allocation, or improvement of the program. This chapter details the characteristics of the study participants, describes the setting of the study agencies, and presents the findings related to the research questions.

Setting

This study focused on employment and training programs administered within Community Action Agencies (CAAs) in Pennsylvania. Community Action Agencies (CAAs) serve diverse functions in the human services field with the specific mission of addressing poverty. These agencies are eligible for Community Services Block Grant (CSBG) funding. Because they receive CSBG funding, they all have the same requirement to use the Results Oriented Management and Accountability (ROMA) framework. This performance measurement model is unique to CAAs, but consistent across all CAAs. In Pennsylvania, CAAs provide information through the Community Organization Planning and Outcomes System (COPOS) reporting system. I further narrowed my sample to CAAs that provided Employment and Training (E&T) programs in order to keep the types of services provided and the types of information collected consistent.

The interviews included individuals working in four separate programs that provide employment and training activities: Work Ready (WR), Employment Advancement, and Retention Network (EARN), and Workforce Investment Opportunity Act (WIOA) Title I. The fourth program was an employment program unique to one agency that caseworkers referred to by various names. For the purposes of this research, I will refer to it as the Employment Placement Program (EPP). Three out of the seven agencies provided more than one type of E&T program.

Caseworkers in two of those agencies worked in multiple E&T programs. Of the caseworkers for E&T programs, four also worked in programs other than E&T. In four agencies, the caseworkers I interviewed also had some supervisory or management level responsibilities. Below I describe the programs from information provided by the study participants.

1. WR is described by the study participants as an employment and training program that is designed to help participants remove barriers to employment through employment readiness skills, community service, and case management. All seven agencies administered a Work Ready Program. Twenty-one study participants worked in the WR program.
2. EARN clients, according to the study participants, are all referred to the agencies by the local County Assistance Office. The Department of Health and Human Services funds this program, according to study participants. The study participants working in this program indicate that the major focus of the program is to help clients find employment with the assumption that clients have overcome their barriers to employment or the remaining barriers are minimal. There was one agency in the study that provided services through the EARN Program, but all EARN caseworkers that I interviewed referred to this program repeatedly as one program to which WR clients are referred to when they complete WR requirements. There are instances when WR clients find employment because of WR activities and are not referred to EARN.
3. WIOA Title I is funded through the Pennsylvania Department of Labor and Industry. Two study participants are caseworkers in this program from one agency. That agency also provided services through WR and EARN. Another agency provided services through WIOA Title I, but no staff members of that program in that agency were part of this study.

4. EPP was unique to one agency and funded through internal agency resources. This program focused on job placement, specifically, and not job entry skills or barriers. Four study participants worked within this program and their agency provided other E&T services through WR. The study participants only identified the COPOS report as the performance measurement used for this program. This program accepted clients from any referral source as well as walk-ins who met the income eligibility requirement. As a new program, study participants could not describe the program in detail; the sense was that the program was in development and the agency was making major changes in their organization and this program was evolving to meet community and client needs.

Agency Descriptions

In order to capture individual perspectives as differentiated from the organizational culture of an agency, I recruited participants from seven agencies. The goal was to have a cross section of urban and rural as well as large and small. The reality of the employment and training programs within the agencies included in this study was that the employment and training components of the agencies were small regardless of the size of the agency. The least number of staff members working in an employment and training program at any agency was two and the largest number of E&T employees identified at an agency was six. All of the agencies had more than one office location in their service area. All of the agencies had E&T staff working out of multiple office locations. Three of the seven agencies worked in multiple counties.

Based on the United States Census Bureau classification of counties as either mostly rural or mostly urban, five agencies represented mostly rural areas and two agencies represented mostly urban areas. I also considered revenue as reported through 2015 IRS 990s filed by six of the seven agencies. Agency revenue includes revenue across all programs administered by the agency. The

average revenue for the six agencies for which the 2015 IRS 990 was available was \$6,728,000. The mean revenue for rural agencies was \$6,234,000. The average revenue for urban agencies was \$7,717,000. Specific dollar amounts are not included by agency so that agencies and study participants are not identifiable.

Although I intended to include only nonprofits in this study, one agency was a public CAA. The distinction for the purposes of this research is minimal because the agency has to follow the same performance measurement standards as the nonprofits for the programs discussed. I was unable to determine size based on revenue for this agency. There are some structural differences between a public CAA and private nonprofit CAA involving governance: a public CAA is under the direction of the county commissioners where a private nonprofit is under a Board of Directors and maintains its own 501 (c) 3 status. There was no sense from the staff at the public agency that not being a NPO affected the topics of discussion.

Table 1
Agency Descriptions

Factor	<i>Agencies</i>	
	<i>N=7</i>	percentage
Urban or Rural ^a		
Serving Counties mostly urban	2	29%
Serving Counties mostly rural	5	71%
Size by Revenue ^b		
Not available	1	14%
\$0 to \$2,000,000	1	14%
\$2,000,001 - \$5,000,000	2	29%
\$5,000,000 – \$10,000,000	1	14%
\$10,000,001 and greater	2	29%
Employment and Training Programs		
Work Ready	7	100%
Work Force Investment Opportunity Act	2	29%
EARN	1	14%
EPP	1	14%

^aBased on U.S. Census Bureau classification. ^bBased on 2015 IRS Form 990 retrieved through www.guidestar.org.

Study Participant Characteristics

Of the 28 participants in this study, 21 were caseworkers and five of those shared responsibilities between direct service provision and management. Seven study participants were in management and did not work directly with clients. Those seven included three agency executives, three program supervisors, and one manager of data systems and reporting. Two study participants cut short their interviews. One person declined to answer questions ten minutes into the interview and requested that I provide the questions by email. That individual's responses were brief and largely non-responsive because this employee was new to the agency and the programs. Another interview was cut short after 22 minutes because the participant was called to attend to a client situation and then was not available to complete the interview. This study participant was included in the coding process although some questions were not answered and some demographic information was incomplete. One staff member worked closely with the E&T programs but was in fact an Adult Literacy instructor who worked across several programs. Since the agency considered her part of the E&T staff her interview is included in the data.

Demographic information about the study participants is limited to gender, race, ethnicity, education, and experience. As shown in Table 2, Study Participant Demographics, 22 of the participants are women (78%); five out of the seven participants in management positions (71%) were women. There are five (18%) participants who identify as African American and one (14%) manager is African American.

Table 2
Participant Demographics

Factor	Total Sample <i>n</i> =28		Caseworkers <i>n</i> =21		Management Only <i>n</i> =7	
	Number	Percentage	Number	Percentage	Number	Percentage
Gender						
Man	6	21%	4	19%	2	29%
Woman	22	78%	17	81%	5	71%
Race/ethnicity						
African American	5	18%	4	19%	1	14%
Caucasian	23	82%	17	81%	6	85%

Study participants were also asked about their education and experience as shown in Table 3. Among the 27 of the 28 participants who provided information on their education, four of 27 (15%) had a high school diploma, three (11%) participants indicated some college but no degree. Three (11%) participants indicated they had achieved an associate’s degree. Eleven (39%) participants held a four-year college degree and three (11%) had earned a graduate degree: One worked in management and the other two were caseworkers. Although the specific degree of the participants was not a question I asked, three participants offered that their degree was unrelated to social work or human services. Of the seven participants employed by the agencies in a management position, one had a high school diploma, five had a bachelor’s degree, and one had a master’s degree. Sixty percent of participants had 10 or more years in human services and in their current position.

Table 3
Participant Experience and Education

Factor	Total Sample n=28		Caseworkers n=21		Management Only n=7	
	Number	Percentage	Number	Percentage	Number	Percentage
Education^a						
High school	4	15%	3	14%	1	14%
Some College	3	11%	3	14%	0	-
Associates Degree	3	11%	3	14%	0	-
4- year degree	14	52%	9	42%	4	57%
Graduate Degree	3	11%	2	09%	2	28%
Years with Current Agency^b						
< one year	8	30%	7	21%	1	20%
1 to 5 years	5	19%	4	19%	1	20%
5 to 10 years	4	15%	4	19%	0	-
10 to 20 years	5	19%	3	14%	2	40%
>20 years	4	15%	3	14%	1	20%
Years in Human Services Field^b						
0 to 10 years	8	30%	6	29%	2	40%
10-20 years	7	27%	6	29%	1	20%
> 20 years	11	42%	9	43%	2	40%

^a The total number of responses for education was 27. ^b The total number of responses for years with Current Agency and Years in Human Services Field was 26.

Findings

Analysis of the data from the interviews produced four key findings concerning employee's' perceptions on the impact of performance measurement. Each is outlined below and then discussed in greater detail.

1. The first finding relates to how study participants actually implement the requirements of performance measurement. I placed this finding first to allow readers to understand how study participants described the performance measurement process. Performance measurement data collection is intertwined with orientation and compliance. Although study participants collect large amounts of data, study participants believe that only limited data is being used to determine program effectiveness by funders. Furthermore, the

information that is being used for performance measurement in this study does not reflect the purpose or goal of the program.

2. The second findings suggest that there are *internal and external influences in performance measurement that affect its reliability and validity* as a tool for program improvement or monitoring. There are differences in how study participants are collecting data and the situations in which study participants collect data that are not considered in performance measurements.
 - a) Even though data collected and reported by agencies is the same, internal systems used to collect and report information is different in each agency.
 - b) There are circumstances and actors outside the agency's control that affect organizational performance which are not accounted for in determining program effectiveness.
3. The third finding highlights that it is not just time as a resource that is a concern in performance measurement, but it is also training resources. The process of data collection is consuming resources that affect program delivery and quality.
4. The fourth finding reflects that existing communications do not support performance measurement as a means for program improvement. The findings are presented as follows:

Finding One: Disconnect Between Performance Measurement and Mission

Performance measurement data collection is intertwined with orientation and compliance. Although a large amount of information is being collected at the beginning of programs, this information is not being used to show client progress. The data being collected during and at the end of the program does not reflect the purpose of the program.

I asked participants about their experience with data collection, what types of data they collected, and difficulty collecting that data. This allowed the participants to describe performance measurement as part of their general work experiences. Caseworkers collect information used for performance measurement at the same time as they collect information for orientation, assessment, eligibility, and other agency purposes. Study participants did not make a distinction between these purposes when they described collecting information from clients, making it difficult to separate the process of data collection from performance measurement and data collection for other purposes. There was no sense that the caseworkers questioned the need for all of the information; they collect it as part of their job.

Caseworkers describe data collection as an ongoing process. When a client first comes into the program, the caseworker does an intake. Some caseworkers referred to intake as an orientation. During orientation, the caseworker collects data to determine a client's eligibility for services, both employment and training (E&T) and other services provided by the agency. When a client was referred for E&T services by another agency or program, the caseworker still had to verify information through an intake process. Caseworkers collected information on the family members as well as the program participant.

Collectively, participants in the study described the data they collected using over 150 phrases to describe the information collected from the clients. Client data elements could be grouped into three different categories of related information: demographic data, assessment data, and compliance data. Demographics was a generic term used by caseworkers to include elements such as name, address, age, family type, family size, insurance, income, housing, gender, race,

and veteran status. The majority of data including demographics and assessment were collected at the time of entry into a program, also referred to as orientation.

Client assessment data gathered by participants in this study as part of their work included both qualitative data and quantitative data such as information on work history, personality types, interests and skills, education level, mental and physical health information, mainstream services being received, childcare needs, transportation needs, and personal expenses. Other personal history information gathered from clients relate to drug addiction involvement, pregnancy, learning disabilities, and legal issues. All the caseworkers referred to assessment data as a means of identifying barriers to employment.

In the Work Ready (WR) program, the goal was not for the client to obtain employment, but to overcome barriers. In other programs, the assessment data was used to assist the client in finding employment that met the clients' skills and abilities. As one caseworker put it, "we're in every aspect of their life because our job is to make them not only be able to obtain employment but to retain employment. Okay, but basically if you don't have a place to live, you don't have food, don't have a car, you're not going to get a job and keep it." In the WR program, these assessments are used to determine barriers to employment. This information digs into the personal life of the client even more deeply than other E&T program assessments. Clients are asked about criminal history, childcare needs, status with probation and parole, work history, driver's license, whether they have any fines, their mental health and that of family members, their disabilities and those of family members, whether they or their children have learning disabilities, if they are in drug or alcohol rehab, and if they have domestic violence concerns.

The study participants explained that once clients were in an E&T program, the caseworkers collected compliance data primarily to document time and attendance as their clients

moved through the program. Some of clients had placements outside of the agencies in community service, some attended school or training programs, and some clients did find employment. Clients need to present to the caseworker documentation such as paystubs, medical information, and birth certificates; the documentation is not just for the client, but also for all members of the household. Study participants used compliance data to document program participation by the clients. Compliance included the client's time in the program, activity, attendance, and caseworkers used the phrase "time and attendance" to reference this type of compliance data. Work Ready requires client data on time and attendance: time spent in allowable activities and meeting the required attendance hours. The EARN program requires information about work hours.

Data collection is an on-going process. The process of orienting new clients to a program is how caseworkers learn who the client is as a person. What are the client's needs, why do they have those needs, and how can the agency help the client? Caseworkers use this assessment data to determine the best course of action, or appropriate goals, for a client to work to accomplish. As one caseworker participant said, "we collect everything." Agency staff members at intake collect both basic demographic and assessment data, and they have to document all of this information.

Although caseworkers told me that they collect information on clients' barriers to employment at intake, these data are not the performance measurements sought by the funders of the Work Ready program. For funders, a successful outcome for Work Ready clients is four weeks of meeting time and attendance goals. Although not the primary goal of the WR program, as in the EARN program, obtaining a job is considered a successful outcome. Documentation for this is attendance sheets, documentation of excused absences, and paystubs from employers. If the client is in community service, evidence of attendance at the community service site is required. If

the client is in training or a higher education program, class attendance documentation is required. In essence, a record of client activities during the time they are in the program is about compliance with required hours and allowable activities, but not about overcoming barriers to employment that were identified at intake. All the agencies that were involved in this study are also Community Action Agencies (CAAs). As a CAA, the caseworkers who participated in this study reported that they did collect outcome data on clients who overcame barriers, such as obtaining a driver's license, securing reliable childcare, or earning a high school equivalency diploma, which in Pennsylvania is called a Commonwealth Secondary School Diploma (CSSD).

I asked a follow-up question of the participants in the study about how the information they collected from the clients was used in terms of service delivery. I wanted to explore how the performance measurement data was used to improve how the programs serve clients. I wanted to elicit from the study participants whether they felt that performance measurement data were used for program improvement or whether they felt the data collection was a matter of accountability and monitoring. I asked this follow-up question of 23 study participants; 22 responses (96%) indicated performance measurement was used predominantly for compliance or monitoring though three responses (13%) indicated that the information was used for program improvement. When I asked to provide an example of how the program delivery was improved, the study participants were stumped. One study participants indicated that the data informed the agency that clients were more successful when the different sections of Commonwealth Secondary School Diploma test were administered at different times instead of having clients take the test all at one time. Since studying and testing for the high school equivalency diploma are not allowable activities under state funded Employment and Training Programs, this was an agency level use of data and corresponding change rather than one based on performance management accountability

reporting. This indicates the agency was using the data internally and that impacted the clients in the E&T programs.

There was a sense that the study participants hoped or assumed the client data they collected were being used to improve the program at the funding level, but in my interviews, they provided no clear indication that they saw those changes. As one manager put it, “You would think and hope that they (the funder) would.” One participant indicated that they thought the goals of the program were changed because no one was meeting them. Another study participant indicated that at one point one program allowed for outreach to clients before intake, but agencies were no longer doing this activity. It is not clear whether that was a program change due to funding cuts or an effort to improve the program; it is notable because of the intense efforts the caseworkers put on client engagement. Client engagement and outreach are downplayed by the funding sources, according to the study participants, even though caseworkers are still working to build relationships with clients as part of improving their program effectiveness.

To explore more deeply into how the study participants felt about the potential the data they collected could be used, I asked participants the question, “From your perspective, what would it take to make the information you collect more meaningful?” I wanted to get a broader picture of how the caseworkers felt about the performance measurement process. Three (1%) study participants did not have a response for the question. The most common response (13 out of 24 responses or 54%) was that the E&T programs were not collecting the right data. Typical statements I heard were “Sometimes the good data doesn't always capture everything about a person because you almost need more information,” and “There is no real place to put what really happened in the system.” I was surprised by a sentiment shared with me by two caseworkers who said, “As much as I complain about ORS, I actually kind of wish the CWDS was more like ORS

because I feel like that data collection actually tells us something.” The second caseworker had a similar comment, “if it (CWDS) collected more data like ORS does...where you can actually put in more goals, I think it might make it a little bit more meaningful.” CWDS is the Commonwealth Workforce Development System data base system used for the E&T programs and ORS is the data base system used to collect data for the agency’s CSBG reports. The Work Ready programs are only interested in time, activity, and attendance of clients despite the WR goal of overcoming barriers to employment. By contrast, the agency level database system used to collect data for the CSBG reporting collects information on clients’ progress in obtaining transportation, education, childcare, and tracking other positive changes that alleviate barriers to employment. A participant who is a supervisor and a caseworker stated:

It gets harder to communicate with [the funder] because our performance report doesn’t ask you how many people did not have childcare before they came to Work Ready and how many people got to obtain childcare and were able to get a job. Those are the questions that are good questions the question [*the funder*] is asking is how many people have their hours in this month and what were the projected hours that you're supposed to have for the month.

On the same topic of making performance measurement more meaningful, a high-level manger was adamant that the data collection should take place over a longer period of time and be more extensive. According to this manager, the timeframe that data was collected on a client, 90 days, was not enough time to make sure that employment was stable. This manager also said that more information should be collected to make sure the agency was doing everything they could

for the client. The barrier to more longitudinal and extensive data collection for this manager is resources.

Finding Two: Performance Measurement Processes are Inconsistently Implemented

Study participants described the processes they used to collect performance measurement data within their agencies. Program guidelines and actual program delivery is beyond the scope of this research, but study participants brought up these factors in talking about training and compliance. State funded employment and training programs have extensive guidelines. One caseworker presented to me a program guideline book that was 562 pages in length. Despite the extensive guidelines, as study participants described their experience with performance measurement, variations and inconsistencies were revealed. These differences occurred at the agency level and included the process for collecting data, staff caseload, trainings, and qualifications. External differences included local County Assistance Offices (CAOs) and the clients themselves.

Internal influences on performance measurement processes. Work Ready program guidelines became part of the conversations about data collection since the predominant focus of data collection was on being compliant with program guidelines. Of the 23 study participants whom I asked the follow-up question about how the data was used, all (100%) said that it was for compliance or monitoring. All study participants indicated that monitoring agency and client compliance with program guidelines is the primary reason for data collection. In one agency, a caseworker who also had management responsibilities expressed her frustration that there was no organization to the updates received through email or informal conversations. She essentially had a collection of emails that clarified guidelines, but they were not in any organized fashion such as a Frequently Asked Questions (FAQ) document. This suggests that although program guidelines

were comprehensive, updates and guidance on the guidelines was not consistent or sufficient. The latest program guideline update available to the caseworkers at the time of the interview was dated 2012 and was only 12 pages in length.

Data collection processes vary across agencies. When I asked study participants about their experience with data collection, I got a sense of how extensive the data collection process is. Their descriptions also revealed differences in how they collected and recorded information at each agency. Caseload, staff members assigned to data entry, and the process for data collection all varies across agencies. Time spent in data collection was of specific interest to me, and I asked study participants how much time they spent on data collection.

The WR, EARN, and WIOA programs use the Commonwealth Workforce Development System (CWDS) to record information. All the agencies also had a separate data base system for internal use that was needed to meet CAA funding requirements. The agencies in this study used four different systems for agency wide data collection. They ranged from internally developed systems to nationally available systems used to complete the Community Organizational Planning & Outcome System (COPOS) report. In Pennsylvania, the COPOS report is submitted by each CAA to the Department of Community and Economic Development (DCED) to meet the Department of Health Human Services reporting requirements (Department of General Services, 2011) for Community Service Block Grant (CSBG) funds that are unique to CAAs. Caseworkers were completing data for the employment and training programs requirements as well as data collection for agency wide reports.

Study participants at each agency had a different caseload. All agencies had some degree of shared caseloads among the caseworkers; five supervisors also shared caseworker responsibilities. The number of clients being served varied widely at the time of the study; one

agency had two clients, and one had 50. Data collection, tracking, and documentation requirements varied by clients depending on which program and what phase of the program the client was in. A caseworker could have multiple clients with each in a different phase of the program at one time. Some agencies shared caseloads across different programs; this meant the caseworkers could have clients in several different programs at one time. This made the question of how much time the study participants spend on data collection and data entry difficult for them to answer because it depends on their caseload, the programs the clients were in, and in what phases of the program the clients were in.

The study participants described tasks involved in data collection and entry, and the tasks were divided differently in each agency. One strategy was for the caseworker to do all the data collection and data entry as well as assist with the necessary reporting for their own caseload. Another strategy was for the caseworkers to collect the data but have another staff person complete the data entry. Yet a third strategy was for one caseworker to enter data into the CWDS system and have another staff person enter it into the agency database. One agency used volunteers to do the data entry and another agency had a dedicated staff person assigned to do data entry across multiple agency programs. All seven agencies at which study participants were employed had other staff members, outside of the E&T program, involved at some level of data collection and data entry, from receptionist to management to volunteers. Agencies also used different strategies at different times.

The performance measurement data collection and entry processes at each agency varied so greatly that I could not compare how much time study participants spent doing data collection, either in general or specifically for performance measurement. The variation in time spent on data collection and reporting described by the study participants was striking. In one agency, the

executive director stated, “I have a counselor that’s spending 20 hours a week just doing data entry.” At this agency, the caseload was six – 10 clients. This does not include a full-time staff person who manages the agency database for all the agencies programs, not just the employment and training programs. At this agency, they also spend one full week on the intake process. At another agency when I asked a caseworker how much time she spent on data entry she said “A lot. If I start getting behind I can spend days at a time doing nothing but data input.” This particular caseworker had a caseload of 50 clients and was referring to just time spent putting data into the system, not including time required for the data collection. At another agency the response was “some days more than others, but typically two hours out of a seven-hour day. You spend more time doing paperwork than helping the client.”

The responses to the question about the amount of time devoted to performance measurement was very different across different agencies. At one site a study participant told me, “It doesn't take a lot of time because I keep up with it. Every day I'm working on it. I would probably say, as far as the time, maybe fifteen minutes just to do the hours. But then when I'm working on goals and I'm doing, you know, some case notes, that's depending on the day and what's going on -, I don't even know if I can put a number on it, really, because I'm doing it and it is bits and pieces. It's constant throughout the day, a constant process.” At another agency a caseworker stated that she spent on case notes for each client everyday “I wouldn't say more than fifteen minutes.... yea not even fifteen minutes I'd say once a week to enter attendance to update activities per client.” I received conflicting information from study participants about whether case notes were required or not. One caseworker who entered case notes into CWDS clarified that her supervisor required it but “Overall it's not a regulation. [My supervisor] is basically saying ‘I think we should.’ And I think [the supervisor] is right you know... but it's only because [the

supervisor] thinks so and our agency is saying so, not an outside party.” The differences in the time spent in data collection is also reflected in the manner in which data is collected.

How study participants, who provide services directly to clients, collect data also varies across agencies. The study participants referred to this process interchangeably as “intake” and “orientation”. All the agencies had caseworkers who collected information from clients face-to-face, used information provided by other agencies, and used forms. Caseworkers in five agencies also collect data from clients in a group setting. One agency collects data in public spaces, such as a library, that is easy for clients to access. Four of the seven agencies (57%) also used formal assessments including the Meyers Briggs Type Indicator, TABE testing, and self-developed inventories. The study participants use these more formal assessments to tailor services to meet the client’s needs. However, study participants gave no indication that these assessments were required by the program guidelines. I asked the caseworker that mentioned the Meyers Briggs Type indicator if they had any training on how to administer the inventory, and I was told they had not. As the participants described it, the process of orientation took anywhere from one day to one week, and differed within each agency depending on the caseworker doing the intake. All the agencies described a process that occurred over time.

Qualifications of caseworkers vary. Variations in caseworkers’ education and experience, coupled with the lack of unified training in program delivery, leaves room for a wide range in staff capacity that could affect performance measurement. I asked study participants how long they worked in E&T programs and/or human services and how long they worked at their particular agency. The length of time at the agency for all study participants ranged from one month to over 25 years with 60% of participants with 10 or more years of experience in their current jobs. The study participants with responsibilities in both management and casework all

had over five years of experience in their current agency and their time in human services ranged from seven to 30 years. The caseworkers also had a wide range in experiences. There were five study participants (24%) working directly with clients who had less than six months experience in their current position, but those five study participants had a total of 77 years' experience in the human services field. On average, participants have 15.4 years of experience within a range of human services. Prior experience of caseworkers included foster care, employment and training at another agency, homeless programs, and foster care placement services. As it became apparent to me that the experience of the study participants varied widely, I asked two study participants from two different agencies whether there were any formal training or educational requirements for the position. Both participants indicated that they assumed there was no requirement specified by the program guidelines. There was not a large enough sample size of participants in this study to make any interpretations about how experience may affect perceptions of performance measurement activities.

The educational level of the study participants also varied. In management, four of seven (57%) had a bachelor's degree, two (28%) had a master's degree and one (14%) had a high school diploma. For the five study participants who were in management but also provided direct services to clients, three (60%) had completed their bachelor's degree, one (20%) had an associate's degree, and one (20%) had a high school diploma. The variation at the caseworker level was that eight out of 16 (50%) had a bachelor's degree, one had a master's degree, one had an associated degree and five (31%) had a high school diploma and one is unknown.

Despite all the variations in experience and education among caseworkers across agencies, there was one overwhelming similarity: 17 out of 21 (81%) of the staff members who were caseworkers discussed the importance of building relationships with the clients. Of the four who

did not mention this, two (50%) were not involved in the orientation process. A third stated, “We serve the client above and beyond” when referencing the benchmarks set by the state. This indicated to me that there was a high degree of commitment from the study participants, regardless of differences in education, to help the clients achieve their goals.

External influences limit program effectiveness. Study participants, and not just caseworkers, mentioned factors beyond their control and outside of their agency as influencing the effectiveness of their programs. There were three specific sources of external influences identified by the study participants: County Assistance Offices, the clients themselves, and program guidelines. Those influences were not reflected in the performance measurement. I found no patterns to the responses that indicate that organization culture or staff characteristics contributed to this finding, but it is clear that external factors affect the programs and these impacts are unaccounted for in the performance measurement the funder is using to evaluate the programs.

The first external influence on the effectiveness on the programs in this study was that of County Assistance Offices. The CAO is the referring agency for Work Ready and EARN programs. The CAO has additional resources to help clients with things such as car purchases. The CAO also has the authority to sanction clients, which may result in a loss of benefits if they are noncompliant. Thus, the CAO influences the motivation of the client to meet WR and EARN program requirements. Three agencies provided services in multiple counties and therefore worked with different CAOs. Each one of these agencies described working with the CAOs in each county as influencing how the agency was able to serve the clients. For example, in one county, the CAO accepted online training courses, but those same courses were not accepted by

the CAO in another county with which the agency worked. The same issue surfaced in car allowances; they were easier to get in some counties than in others.

One responsibility of the CAO is to hold Direct Service Team (DST) meetings. DST meetings were mentioned by eight (28%) of study participants as an important part of the E&T programs. Study participants mentioned these meetings in the context of reporting, training, and feedback. The DST meetings involved discussing clients and their challenges and successes. The meetings were led by the CAO for the EARN, WIOA, and WR programs. Attendance at the meetings varied by county but included agencies that provided support services to clients such as Child Care Information Services, which provides subsidized childcare for participants who are working up to 20 hours a week, other E&T program caseworkers, and education providers. Since there may be more than one agency involved in administering the EARN, WR and WIOA training programs in a county, other employment and training agencies were also included in the meetings. The frequency of these meetings also varied across counties from once a month to once a week.

DST was described by one caseworker as a time when she can “discuss any problems, any issues, positive or negative with the client, and the best way to help them out; and it is feedback.” DST meetings are a venue where caseworkers can ask questions, seek advice, and get feedback and validation. Although DST meetings are required and documented, I sensed that the documentation and discussion was less formal than a technical assistance request or monitoring visit. The caseworkers participating in this study also consider DST meetings as a type of training. Participants viewed DST meetings as an avenue to get information and advice on how to handle client situations within program guidelines.

Study participants made only positive comments about the DST meetings. With the variation across counties in how DST meetings were conducted and the informal training process that occurred at these meetings, the local DST meetings serve as an important source of support and guidance for caseworkers on how they handle cases and how programs are administered at the local level. That caseworkers experience with DST's varies also contributes to inconsistencies in their training and communication around performance measurement.

The second external influence on the effectiveness on the programs delivered by study participants are the clients themselves. Participants voiced considerable frustration that the performance of caseworkers and the agencies depends on decisions made by clients. For instance, the most predominant factor used to judge the performance of agencies administering the WR program is the participation rate of the client. Study participants referred to this as "time and attendance." If a client does not participate according to program requirements such as failing to show up for class, work, or community service for the required hours each week, the agency must count this as a "negative termination" of the client's case. This does not consider the client's situation or the client's role in non-participation in the program. Even caseworkers who indicated strong empathy for clients felt that client decisions affected how the state judged their agency's effectiveness. One caseworker with four years of experience explains it this way:

[The client] doesn't come in because 'I'm angry I don't want to be here' or the biggest one I run into is drugs. Absolute substance abuse. I have had ones that, honestly, we have had them come in high, which you give them credit at least they showed up, but they come in high. They come in drunk or there are criminal acts and they're running from the police. You know a

bunch of different things going on and unfortunately you have some that they're still stuck in that 'I don't understand why I need to show up.'

Another caseworker talked about a client who had been through the program seven times. The agency had a negative outcome recorded for each of those times.

Study participants also shared positive stories of clients. One young mother with several children was admired by her caseworker for her perseverance. This client had to walk two miles with her children to get them to day care before she could go to WR. Another caseworker shared this story:

There was still a lady that would walk here. From the other side of town. No matter what the weather was like it even when we had temperatures of negative twenty. She would walk here and make it your own time every single morning. You know sometimes I was like what I'd be able to do that if I had to walk a mile in twenty below when the wind and the snow would I would I do that. I question that a lot and so whenever individuals like that succeed at the littlest things: I think that's a big deal and should be looked at.

Client characteristics are more than core demographics and information in a file for the caseworkers. They recognize the differences in some clients 'motivations and perseverance and believe that those differences should matter to the people evaluating their effectiveness within in the program. They may share these stories at the DST meetings, but there was no sense that these stories of the small successes had any significance at a level that would impact policies and decisions.

Another question I asked study participants to provide insight into client behavior was whether the caseworkers had difficulty getting any performance measurement information needed

from the clients. Twelve of the caseworkers (57%) responded that they did not have difficulty getting the data from the clients, but four of those case workers qualified that by stating that sometimes a client could be difficult. Four caseworkers indicated that they did have difficulty getting the data and attributed that difficulty to the lengthy process, lack of client trust, not having the right documentation, or the inability of the client to provide the information. At three agencies, caseworkers mentioned that clients were often high or on methadone and it was difficult to work with them. One experienced caseworker stated that “clients become highly indignant about us delving into their personal business,” but she seemed to take the reaction of the clients in stride as just part of doing the work.

The third external influence on the effectiveness of the E&T programs is the program guidelines. As one caseworker put it “the [program] guidelines are ... cumbersome.” Client eligibility was not determined by the agencies for all the employment and training programs. For WR and EARN, eligibility was determined by the CAOs. But in the WIOA programs and support programs like adult literacy, eligibility was viewed as one factor in program guidelines that restricted the help a client received. A prime example is the high school equivalency program. It was beyond the scope of this research to dig into the details of how eligibility was determined, but study participants explained that there are age restrictions and allowable activities that precluded some clients who needed a diploma from working on that goal as part of their participation in the E&T program. Study participants also mentioned the restrictive program guidelines and reservedly said that they had to be “creative” or help clients “off the books.” Some study participants who are caseworkers reported helping clients without the agency knowing. (When study participants shared this information, they sought additional assurance from me that their identities would remain confidential).

Finding Three: Performance Measurement Takes Away Resources from Programs

The time spent on meeting performance measurement requirements is of interest to me because it is an aspect of the impact that performance measurement has on the services provided to clients. Based on what I heard from my study participants, time spent on performance measurement takes away from time needed to do the actual work of serving the clients. As noted in the first finding of this study, performance measurement is intertwined with other program processes, so time spent on data collection is not specific to performance measurement: that time includes assessment and compliance. However, in my conversations with study participants about training for performance measurement, there emerged another resource that data collection was using: training time. The state prioritizes correct data entry, so understanding the CWDS system takes priority over training caseworkers to better serve clients. Study participants referred to “time and attendance” and getting the coding correct as the main purpose of training. Though not a specific question in the interviews, some study participants also mentioned training on compliance with program guidelines. What appeared to be missing was training on actual program delivery to the clients; the focus was on getting the data entry correct according to guidelines but not on how to serve the client. I asked a person in upper level management during one of the last interviews for clarification on the trainings emphasizing correct data entry as opposed to trainings on how to help clients overcome barriers. The response to me was, “I agree with you 100%,” and this manager further stated that “they [the caseworkers] have to make sure they put this person in a proper category and so forth that’s what the state’s more concerned with.” Training resources spent on data collection becomes a concern because there are no required qualifications to be a caseworker in this field.

Participant responses indicated that there was some effort at the agency level to improve caseworker performance. An example of agency level training designed to improve service delivery was the Strength Based Family Worker (SFW) credentialing program. One agency was sending all their caseworkers to this extensive program that works with client engagement (Temple, 2018). Another participant, a caseworker with extensive experience, reflected that the emphasis on data collection was not always the case: “They wanted you to go and be part of part of trainings where they would try to help the good people do better case management. Like I said, those were back in the days when it was all about engaging the client.” A participant in management explained, “We are spending all of our time on the paperwork and trying to make sure all the ‘Is’ are dotted and ‘Ts’ are crossed so that we’re spending less time to work with the client.” Yet another caseworker participant stated, “We could do some more workshops on probably better helping clients with their needs. It's almost like, you’ve got your client, work on it. Do the best that you can. If you need any help, get a hold of the CAO as to what you can and what you can’t do.”

Study participants indicated that training was inconsistent in terms of content. I asked study participants what type of training they received for collecting and using the performance measurement data they gathered from the clients, and what type of training they wished they received for this work. The funder, which is the state, provided E&T program training. Agencies provided trainings, as well. There were four types of training modalities: online, in-person, email, and conference calls. According to the study participants, the primary objective of the trainings by the state and the agencies was correct data entry. No participant mentioned any training that would assist caseworkers in using data to improve to improve programming.

Training is inconsistent, too, in terms of frequency, according to study participants. Trainings provided by the state included a two-day Commonwealth Workforce Development System (CWDS) training where staff traveled to the state capitol. This training is intended for new staff members, but it is not offered on a regular basis. Though study participants' experiences varied, they consistently indicated that this training occurred one to two times a year but not predictably. One study participant said she missed the training because she did not know about it until it had passed. Another caseworker mentioned that she cancelled her vacation to attend the training but then the state cancelled the training. Not all E&T staff had received this CWDS training. It was not required: some staff who entered data had not had the training. Eight (30%) of study participants said other people in the office trained them and two (7%) had received the training through previous employment. Four (15%) of study participants (a manager and three caseworkers) said they trained themselves. Both caseworkers and managers mentioned they had contact persons with the state funding agency whom they could email if they had questions.

I observed from participant responses that non-standardized training occurred across all the agencies represented by the participants in this study. In five of the seven (71%) agencies, study participants taught themselves data collection or data entry for performance measurements. In six of the seven agencies (86%), other staff members trained study participants. In the one agency where there were no study participants trained by other staff members, 66% of the agencies E&T staff had been at the agency for over 20 years.

Training opportunities were not only unpredictable, but they were increasingly rare. In reference to trainings, one participant with eight years of experience, a caseworker with management responsibilities, recalled, "In the beginning it [training] used to be a lot. They would

offer several trainings a month. Then you would go to at least one a month sometimes more”.

That same study participant shared her disappointment in the current situation:

We were supposed to have regional Work Ready meetings several months back. And at that, they were supposed to give us an updated manual, because there has not been an updated manual in several years. They were supposed to let us know of how the program was changing, how they were doing away with the phases and any changes with code we were allowed to use or not allowed to use. But they had postponed those, and still haven't rescheduled them so, nobody has yet received the new updated manual. It would just be nice to find out any updates and any new information, plus anything that I should be doing that I'm not regarding entering information into the system.

It is important to note that this participant moved into a management position and, in her own words, was “self-taught” but did receive guidance from the “lady from the state” who did a visit and let her know what she should be doing.

How performance measurement is used for resource allocation is unclear. One of the purposes of performance measurement is resource allocation. This impacts service delivery at the programmatic level by removing funding from ineffective programs and allocating it to more effective programs. This study found that the activities involved in performance measurement took time and training resources at the caseworker level. Resource allocation also happens at the organizational level by allocating funding according to the performance of the agency. To understand the staff members perception of how resource allocation occurred at the organizational level, I asked the question of how their funding was affected. There were nine (33%) of study

participants that were unsure if funding was affected by performance measurement. This included two participants who were in management positions. Of those that were unsure, they all speculated that it could or that it may affect future funding, but they were not certain. There were 11 (41%) that said that performance measurement did affect funding. These responses varied as to how it affected funding. One manager said “yeah, if we don’t meet our goals it will affect funding”. Four participants shared that the agency used the information to go after other grants. There were comments that suggested that funders had put other agencies on probation or lost their funding due to not being compliant. One manager shared his view that “it is very important because our funding depends on it. You know we have to show these outcomes so if we don’t put that information in right it could affect us, it could affect our funding in the future.” This comment highlights the disjunction of performance measurement. It is viewed as important; to get the information entered correctly to preserve funding. This suggests compliance is more important than actual effectiveness. Actual program effectiveness was given in one response; “If your program’s not successful then you have less of chance to get the money for the next program year. That’s where it really matters, for the funding purposes.”

There were six, (22%) of the study participants that indicated that funding was not affected by performance measurement. Among all of the responses, there were four study participants that felt politics affected funding; who was in office or how information was presented to legislators had more impact on funding than performance measurement. There were two study participants who had heard of an agency losing funding due to compliance issues. It is unclear if they were speaking about the same agency or if this was in fact true. There were five study participants, each from a different agency, which said their funding had not changed; in one

case the study participant specifically stated for the seven years of their involvement in the program the funding had not changed.

The study participants had not experienced funding changes due to performance measurement, but 41% felt that it could. For these caseworkers, performance measurement is more about compliance than performance. It is also a future expectation; if organizations do not meet requirements or show their effectiveness through metrics, their agency could lose funding and miss opportunities for future contracts or grants. While I was unable to determine with any degree of confidence how programs were impacted through resource allocation due to performance measurement, the variances in the responses provided some insight into the issues that occur with communication.

Finding Four: Patterns of Communicating Performance Measurement Results Do Not Reflect a Feedback Loop for Program Improvement

I asked study participants questions about feedback and reporting in an effort to understand the communication patterns concerning performance measurement. Feedback questions led to some conflicting remarks. In one instance, the caseworker said, “If you are missing something they let you know. If you don’t do something right they let you know.” However, that same caseworker reported not getting any feedback. Although 19 of 28 (68%) study participants responded that they did receive feedback, some qualified this by saying “recently,” or “only if we do it wrong,” or “no news is good news.” Eight participants indicated that they did not receive feedback.

The term “feedback” seemed to have different meanings to different people. Study participants described feedback in terms of compliance and proper data entry in 18 (64%) responses. However, 12 of the 21 (57%) caseworkers gave descriptions of feedback they received

for positive work with the clients. Five of those 12 (42%) participants described the feedback as coming directly from the clients. In two of those five (40%) instances, the caseworker helped the client but not within prescribed protocols; supervisors were unaware of the clients successfully getting support because the caseworkers felt they would get in trouble if the agency knew. This occurred in two different agencies, indicating it may not be an isolated situation.

Considering the totality of participants' comments about feedback, not only as a response to the specific interview question about this but also as it was mentioned throughout the interviews, 11 (39%) participants were dissatisfied with the quality of the feedback they were receiving and four (14%) specifically mentioned the poor timeliness of the feedback. A total of eight or only 28% were satisfied, and in nine (32%) responses there was no positive or negative connotation to the responses or the responses were conflicting.

There were differences in participant comments by position and agency. Study participants in a management-only position represented four of the seven agencies in the study. Three of seven participants (43%) in a management only position indicated that they gave feedback to staff. However, in two of the three agencies those managers represented, at least one caseworker felt as if they did not receive feedback. In one agency, all the caseworkers said they received no feedback from management, but the participants in management said they gave feedback to caseworkers. In a different agency, all the caseworkers reported receiving feedback from their agency's managers. The agency with this disconnect had their E&T programs and caseworkers at a location away from the management and central office, but there was no indication that the other offices had similar communication gaps due to physical location. In three of the seven agencies where no study participants had a management role, I was unable to find a pattern in feedback related to physical isolation of staff.

I also asked participants about the type of reports that the agency generated from the information the caseworkers collect from the clients. The purpose of this question was to understand how funders and agencies use the information in terms of program delivery to clients. The responses fell into a pattern that reflects a flow of information from the caseworkers to the managers, but no information getting back to the caseworkers. The reports were in three types: reports for internal use, reports required of CAAs, and reports required by program funders. Two (7%) study participants indicated that they collected data and produced reports for their own use. Interestingly, one of these people did not have a management role. The reports to the funders were through the CWDS system, and the primary item that study participants described reporting on this system was clients' time and attendance in approved activities for the purpose of compliance.

Of the 21 study participants who were caseworkers, 10 (48%) said they did not know what reports were generated from the information they collected from clients. Of the 11 (52%) caseworkers who did know about the reports generated from client data, seven of them also had management responsibilities. None of the 16 study participants without management responsibilities saw the reports. All of the seven (100%) study participants who were managers without caseworker responsibilities knew about the types of reports generated from client data. Of all of the participants who were caseworkers, including those with management duties, 13 (62%) did not see the reports. Participant responses indicated an overall sense that caseworkers see their performance measurement reporting responsibilities as entering client data correctly; by entering information into the data base systems they were engaging in the reporting process. This was supported by two (7%) of study participants who also indicated that they pulled reports to verify that data was entered correctly.

Of the 26 study participants whom I asked about the type of input they had on the data collection process, nearly half (12, 46%) indicated that they had no input into how the data was collected or what information was collected. Of the 14 (53%) participants who said they did have input into the data collection process, 13 perceived having input at the agency level. Only three participants, two of whom had management responsibilities, felt they had input on data collection processes at the funding agency level. Of these three participants, the one caseworker who was not involved in management said the funder was asking what could be done differently with the E-data system. This staff member worked in multiple programs and the referenced system was not specific to E&T programs. One manager who did not have direct contact with clients, indicated having had input during the development of the CCA state reporting system: this system is not an E&T reporting system but is specific to CAAs. In addition, one manager served on a regional committee that provided suggestions on the data collection process within the E&T programs.

All but one (83%) of the study participants who have management responsibilities responded to the question about types of input they have about collecting data, took the perspective of getting information that they need for reports from caseworkers in their remarks. The sole management study participant who answered from the perspective of getting information from the client said there was no difficulty but qualified that statement by saying as long as the client was willing to give it. Another participant response indicated that the staff was good with working with clients. On the other hand, two study participants in management said they had some difficulty getting information needed for reports. One manager referred to the "sheer volume" of information required, and the other indicated that getting information was a problem in the past but was getting better. One participant in management said that sometimes they had

difficulty getting information for reports and needed to press the staff to provide the information in a timely manner.

The key report from the perspective of managers was the COPOS report required of CSBG eligible organizations. As one supervisor describes it, “Everything goes into that major report. So no matter where the funding comes from, CSBG, Community Service Block Grant, our major funding here, everything that we do is considered to be the leverage against those funds we receive.” This report collects more information than required by the E&T programs as far as client outcomes. It includes numbers of clients that receive a job, obtain childcare, get a driver’s license, increase employment skills, maintain a job for 90 days, or obtain their GED. Despite the objective of the WR Program to help clients remove barriers to employment, there was no indication from the study participants that any information on the removal of barriers was reported through the CWDS system. For example, multiple caseworkers mentioned that a lack of a high school diploma or equivalent is a barrier to employment. However, WR program guidelines prevent them from addressing it as part of the program activities. When caseworkers help clients get their high school diploma or equivalent, the agency cannot claim this as a successful client outcome in the E&T program. Therefore, the funder does not consider this an agency or client success. In other words, it does not count. For the E&T funding agency, this means nothing in terms of measuring the agency’s successful outcomes.

Summary

My purpose for conducting this study was to understand the perceptions among caseworkers and agency managers of performance measurement. If the intent of performance measurement is to provide information on the effectiveness of programs, then it is important to understand the practices and views of the individuals doing the work and implementing the

performance measurement. The findings of this study show that understanding the impact of performance measurement on direct service provision and agencies is complicated. Caseworkers intertwine performance measurement with the client orientation, assessment, and other compliance measures. Data collection is extensive, but important information is omitted from performance measurement due to internal and external factors that affect the programs. Even the training and staff requirements were inconsistent.

One remarkably consistent pattern in the data from the study participants is that all of the caseworkers presented an impressive attitude of care and concern for their clients; they were going to do what they needed to do to help their clients pursue their goals. Aside from the laudable commitment of these human service professionals to their clients, the findings of this study reflect a host of inconsistencies in performance measurement practices. These include how data are collected and reported, variations in training for the individuals who do this work, how employees of these community agencies perceived feedback, and input to the performance measurement process. The finding that training resources are used for data collection and entry instead of training caseworkers on providing services makes these caseworkers' characteristics important. We should not lose sight of the idea that the human services field is about helping people. If trainings for staff members to provide quality services are no longer a priority and education and experience in providing quality services are not a requirement, what process is there to determine if a staff member is capable of providing quality services to the client? The diversion of training resources to correct data collection and input from providing quality services means that there is no mechanism to ensure the caseworkers providing direct services to the client have the knowledge and skills to be effective.

Researchers consider performance measurement as a means for both accountability and program improvement (Carnochan et.al, 2014; LeRoux & Wright, 2010; Thomson, 2010). One of the concerns that researchers have is that performance measurement creates organizational cultures wherein caseworkers lack the flexibility that is needed to improve programs and organizations (Murphy, 2007). Research in the field of psychology suggests that tight monitoring systems that lead to lack of trust result in less innovation (Amabile, 1979) and less willingness to become part of effective organizational functioning (Organ, 1988). This study lends credence to those concerns, as the findings suggest little indication that performance measurement is used for program improvement; it is primarily used for compliance.

The patterns of communication surrounding performance measurement also indicate that there is little chance for performance measurement data to be used for organizational growth because the results of the information are not getting back to the caseworkers, at least in this study's sample. Looking at the perspective of the study participants, there is no indication that the information is being analyzed at all.

CHAPTER FIVE

DISCUSSION AND CONCLUSIONS

In this study, I explored how the staff members in nonprofit organizations' (NPOs) perceive the impact of performance measurements on their work, the training they get to gather and use the data, and communication about the results of these efforts. To do this, I explored how human services managers and caseworkers perceive the performance measurement process through their experience. For the purposes of this study, I defined performance measurement as the general activities used to collect and report information about a program's effectiveness with the intent of using the information for monitoring, resource allocation, or improvement of the program. I focused on three research questions that reflect elements of performance measurement:

1. What is the impact of performance measurement on service delivery and organizational performance relative to the mission?
2. Do NPO staff members feel they have adequate (a) training and (b) time to collect quality performance measurement data?
3. What type of communication exists concerning performance measurement results in improving the delivery of programs and services?

To gather data, I used qualitative individual interviews of employees in human service agencies that provide employment and training programs, specifically Community Action Agencies (CAAs). I anticipated, based on the literature reviewed and professional experience, that participants in the study would have negative perceptions about the performance measurement process. Instead, I found more nuance in the data. Caseworkers collect data for performance measurement at the same time as they collect data for orientation, assessment, and eligibility. Study participants understood the necessity of collecting the data, but were negative about the

technology and redundancy involved. I felt a sense of resignation, as if the process was a necessary inconvenience. This chapter discusses the findings about each research question and the current literature. I also discuss limitations, directions for future research, and implications for policy and practice before offering my conclusions from this study.

The Impact of Performance Measurement on Service Delivery and Organizational Effectiveness

The issue that inspired this study for me as a human service professional and scholar was the concern that increasing performance measurement requirements for human service nonprofit organizations (NPOs) may harm rather than help these organizations achieve their mission of serving people. In the case of employment and training (E&T) programs in Community Action Agencies (CAAs), the concern for me was that performance measurement practices might interfere with the relationship between clients and caseworkers, and consequently have a negative impact on organizational effectiveness. The findings of this study did not support this concern. As detailed in the findings, 81% of the caseworkers felt that building a relationship with the client was important. I was genuinely impressed with the dedication and caring shown by the study participants for their agencies and their clients throughout the interviews. The caseworkers used the information to tailor their service approach to clients so their efforts met individual client needs.

Caseworkers put effort into how to find a way to help the client despite the regulations. Caseworkers even went beyond what program and agency guidelines allowed, at a perceived risk to their own employment, in their efforts to help their clients. The dedication of these study participants to their clients may not be representative of all caseworkers, but the findings suggest that dedicated caseworkers overcome barriers that performance measurement might impose.

Other researchers support this. Benjamin (2012), Benjamin and Campbell (2015), and Packard (2009) suggest that inter-relational aspects of a human services program are more important than client compliance for organizational effectiveness. The commitment to both the clients and the agency mission by the staff members in this study meant going beyond program guidelines to help clients. This makes performance measurement for program effectiveness difficult because client outcomes captured in the data may not reflect the impact of the program, but instead it may reflect the committed efforts of caseworkers beyond program requirements. These actions of the caseworkers provide insight into the findings of Dicke (2002) that there is skepticism with managers about the validity of information being provided to funders. This is not to suggest that the information is false, but the information that reports require does not capture the true picture of what happens with client caseworker interactions.

Despite the perceived lack of meaning shared by the study participants in this study, they continue to comply with the performance measurement requirements of funders. The underlying idea of institutionalism is that organizations adopt practices because those practices are generally accepted as being part of the established best practices of organizations like themselves (Powell & Colyvas, 2008). This concept provides a framework for understanding how performance measurements are affecting the structures of organizations. The coercive nature of rigid compliance to program guidelines and reporting requirements leads to the decoupling of activity from the formal structure (Meyer & Rowan 1977). The study participants understood the formal requirements of the funders, but they worked around those requirements so their clients could be successful. Study participants told stories of sitting with drug addicted clients for hours as they went through withdrawal, they drove clients to the area food bank after hours, and they provided support for clients to obtain their high school diploma. None of these activities were required by

the guidelines of the employment and training programs their clients were in, but they were required for their clients to succeed.

This decoupling of the activity from the requirements was voiced by the study participants when they expressed their belief that funders are not measuring the right information. Earlier research by Carnochan et al. (2014) identified that performance measurement can create tension between the funders and service providers by requiring an aggregation of data that does not reflect the progress of the clients served by agencies and programs. Considering that the goal of the Work Ready (WR) program is to reduce barriers to employment, it is striking that caseworkers are not actually asked by the funder to report on the reduction of such barriers in performance measurement.

This disconnect is a legitimate concern. The study participants indicated that they understood the importance of data collection. Their frustration or reticence is not with the performance measurement in and of itself but the disconnect between the purpose of the program and the outcomes recorded. This disconnect is an example of the goal displacement described by Merton (1938); when achieving the goal is thwarted by institutional norms. The study participants describe these situations with clients as frustrating. Caseworkers who chose to help the client in ways that are not within program guidelines are experiencing subjective strain (Agnew, 2001). The type of activity these case workers engage in to circumvent the agency policies and procedures demonstrate that there is some degree of social control, but they still demonstrated anxiety about sharing those stories.

The type of interaction the caseworker has with the client also has implications for the health of the caseworker; Krum and Geddes (2000) found that emotional attachment to the client reduced emotional dissonance. In Krum and Geddes (2000) study, the participants voiced their

views that developing a relationship with the client was important to be an effective caseworker, not just for getting the information, they needed from them for performance measurement. That degree of engagement may also mitigate any negative affects of emotional labor. This is supported by the stories about the how caseworkers went above and beyond what program guidelines said they must do and also went beyond what they perceived was allowable within their organization. The caseworkers have built a relationship with the clients that allows them to engage in emotional labor at less cost to themselves, but at the same time they have increased the risk they face for perceived deviant behavior.

This disconnect also has implications for the ability of organizations to show the effectiveness of the program and therefore be accountable to funding agencies. The mission, or purpose, of the largest program represented in this study, WR, is to remove barriers to employment. Within the WR program, the focus of performance measurement was the time and attendance of clients. It is feasible that funders of the WR program determined that an agency was effective and accountable for their funding because they coded activities correctly, according to funder requirements, without actually removing any barriers to employment for clients. This supports the concerns researchers have that there is confusion between performance measurement for program improvement versus accountability requirements (Carnochan et.al, 2014; LeRoux & Wright, 2010; Thomson, 2010). As one executive director stated, using the cliché attributed to Peter Drucker, “what gets measured gets done.” This reflects the frustration of caseworkers who are measuring client activities not related to the purpose of the program but still required in

reporting. Benjamin (2008, 2012) found this pattern in his research on NPOs in general. As one manager among the participants in this study said:

They make changes to the clientele that are referred to us. They make changes to the activities that we can do with them. And the activities that we can do with them don't necessarily match up with what the program design is.

Therefore, we have to find a way to embed those things. The true outcome isn't getting collected.

Dicke (2002) found that staff might misrepresent data in reports by picking from acceptable lists of activities because what reporting what really occurred is not an option. In this study, I heard from participants who found themselves caught in situations where they underreport uncounted activities that benefit clients, or rename activities in data collection in order to meet program guidelines. The intent is not to mislead funders or falsify reports but to account more fully for what the programs are doing in serving clients. However, these practices certainly have the same effect of undermining the ability of funders and agencies to evaluate program effectiveness.

As employees of CAAs, study participants also were collecting and reporting data for funding sources unrelated to employment and training. The overarching mission of CAAs is help clients achieve self-sufficiency. Because the study participants reported data about agency accomplishments as required by the Community Services Block Grant (CSBG) funding CAAs receive, they reported on barriers clients overcame in an entirely different system than E&T funding required. An unexpected response from two study participants was that the CSBG performance measurement system is more relevant to the mission of the E&T programs than the Commonwealth Workforce Development System (CWDS) used for that purpose. As one caseworker in this study put it, "As much as I complain about it [the CSBG data base system], I

actually kind of wish the CWDS was more like it because I feel like that data collection actually tells us something.” This suggests that the caseworkers may find performance measurement meaningful if it relates to the program mission. It also demonstrates that there are ways for performance measurement to reflect the goals of the program.

Further evidence that performance measurement is not being used for program improvement or organizational change was the lack of examples that study participants were able to provide for changes related to information collected and reported. In this study, there was only one instance in which a study participant was able to point to a specific program improvement made through the analysis of client outcomes. This improvement was at the agency level and not a programmatic change or a change requested or recommended by a funder. The change was in how the GED[®] was administered: the agency found that clients were more successful when they sat for the test on separate occasions instead of taking all five sections in one sitting. This example demonstrates that effectiveness can be improved using performance measurement analysis. However, in this instance, the change was limited to agency-level practices. Study participants expressed a hope that funders would use the information they collected for program improvement, but they did not have specific knowledge of this happening. In this study, the caseworkers were the ones delivering the program to clients and it would seem likely that they would know if funders used performance measurement for program improvement or if communication about results and analyses occurred. The pattern in this study is consistent with the research by Liket et al. (2014) who found that nonprofits struggle with program evaluation when funders combine performance measurement with accountability.

Training and Time Commitments to Performance Measurement

My second focus was on whether the study participants felt they had adequate training and time to collect quality performance measurement data. Study participants described quality in terms of coding and putting the information in to the system correctly. As stated above, the program funders required information with client participation and compliance to codes and guidelines. The funders provided training on data entry, although formal training was not viewed as timely or predictable. However, study participants felt they had the support they needed through experienced colleagues coupled with emails and conference calls to the funders. Still, a lack of standardized training raises questions about the reliability of the information across agencies.

How much time staff members spend on performance measurement is a complicated question. In this study, participants' data collection and reporting for performance measurement is intertwined with client orientation and assessment. Most of the data are collected when clients enter a program, leaving minimal data collection about outcomes and accomplishments after services are received. As one manager stated:

Let's face it, it's tough for funding and all that nowadays. It's all data. What we do with the clients is all non-data. But then we have to put it into data.

That's probably the best way of looking at it: working with clients is in a non-data form, but everything that we do has to be data driven.

This is worth noting because there is no indication that the wealth of information collected when a client enters a program is used to determine the effectiveness of the program in overcoming barriers to employment. In fact, the barriers that caseworkers help clients overcome are not the focus of the performance measurement in the WR program.

Performance measurement that does not include the clients' outcomes that are a result of the program does not provide information on program effectiveness. The process affects the clients by diverting time to data collection instead of service at the beginning of the program, but the outcomes data is not being collected with the same rigorousness when clients are leaving the program. In effect, performance measurement about the E&T programs served by the participants in this study is missing data on client outcomes. One criticism of using performance measurement as a tool for program evaluation is the need for baseline data (Dicke, 2002). This study suggests that there may be a wealth of data collected as clients enter human services programs. This baseline data could be used to determine program effectiveness in a manner similar to program evaluation. One aspect of organizational effectiveness is the effectiveness of the services it provides. If performance measurement is not being used to improve programs, then it cannot improve organizational effectiveness in relation to achieving the mission.

Another aspect of organizational effectiveness is the capacity of the staff members. The process of data collection and data entry has become such a priority and the need to get it correct has caused training on data collection and reporting to supplant other types of trainings. Training of any kind was provided in inconsistent ways to caseworkers. And while some agencies provided caseworkers with training on working with clients, it almost seemed an after thought with other study participants. This lack of training is troubling for several reasons. Skills identified with women's jobs are seen as less important: skills required in dealing with clients is one of those skills (Aker, 1990). Aker also posits that jobs that require women's emotional management skills are placed lower on the organizational chart to reinforce a gendered hierarchy (Akers, 1990). Training in interpersonal skills could assist caseworkers in mitigating the negative affects of emotional labor (Krum & Geddes, 2000). By diverting training to coding and data work,

organizations are devaluing the emotional toll on the caseworkers which plays into the paternalistic pattern of both the expectation that women are better at emotional labor and therefore do not need the training and ignoring the real health affects of emotional labor.

Within the organizations in this study, the demographics of the study participants does not intuitively suggest a gendered in organizations the context of predominantly masculine in authority and feminine in the lower positions of power and presumable wage (Akers, 1990; Haines, Deaux, & Lofaro, 2016). The number of study participants in management were limited by design in this study, but 71% of the management staff were women. However, there is still a perception that the work of the caseworker is unskilled, since there is no requirement for education nor is there an expectation that on-the-job training is going to be provided. These perceptions of caseworkers in the employment and training field as being unskilled, leads to the perpetuation of this type of worker as being low paid. The focus on training for data collection and data entry over working with clients can be perceived as reinforcing this narrative.

Communicating Performance Measurement Results

The third focus of this study was on the types of communication that surround performance measurement. The reason for this line of inquiry was to determine if communication exists between funder and agency and agency and staff member in a manner that would support organizational learning and program improvement. One aspect of performance measurement is the ability of the information gleaned from performance measurement to be used for program improvement. The research in the literature on NPOs often is confusing because researchers do not always make the distinction between human service NPOs and other types of NPOs, such as arts and museums (MacIndoe & Barman, 2012). There also is research on NPOs that does not distinguish between the service providers and the program funders or differentiate between NPO

management and the caseworkers (Carachan et. al., 2014; Dicke, 2002). By focusing on one type of program, employment and training, within one type of NPO, Community Action Agency, I was to distinguish perceptions of managers and caseworkers. My findings show that managers and caseworkers have different perceptions about communication of performance measurement. Although that may not be surprising, it reflects one difficulty that agencies and funders have if they want to use performance measurement data for program improvement: feedback to the caseworkers.

My interview questions concerned three aspects of communication: reports, feedback, and input. Study participants actually discussed communication in the context of monitoring, training, and client satisfaction. Some of the information concerning feedback seemed contradictory. Sometimes study participants made remarks about feedback that conflicted with remarks they made in a different part of the interview. In addition, there were staff members in the same agencies whose responses differed widely on whether they received feedback or not. Some were satisfied and as one caseworker phrased it “we are like mushrooms; they keep us in the dark.” Some caseworkers responded that they got a lot of feedback through informal channels, such as conference calls and emails. Still, 18 out of 28 participants (64%) reported that the feedback they received was largely about whether the data was input into the system correctly. The program parameters were very specific about time and attendance, so caseworkers had a sense of whether they were meeting those expectations without feedback from supervisors or from funders. It is clear that the study participants’ perspective is that performance measurement is used primarily for monitoring and control, not for program improvement.

In nonprofit human services organizations there is a potential for multiple feedback loops between multiple stakeholders, funders and clients. Since I asked study participants about

feedback in general and did not distinguish between different possible sources of feedback, the study participants could have been referring to multiple feedback loops in their responses. The most intense positive responses about feedback were given by caseworkers when they shared stories of clients saying thank you when they saw them in the community. One participant shared that she once was recognized with the employee of the month award. In another agency, a program official gave employees compliments outside of formal channels. This study does not provide the framework to look at communication in more depth since the questions were not structured to make the distinctions among types and sources of feedback: I simply asked whether study participants got feedback, and they shared how they felt. Overall, I learned that these caseworkers are satisfied with the feedback they get from their clients and colleagues. Feedback from funders or managers had a negative connotation because, as one study participant explains, “welfare [the county assistance office] really doesn’t report back unless there’s something wrong.” This suggests that feedback was not used to improve programs or increase organizational effectiveness but to ensure compliance.

Gouldner (1954) identified two ideal types of management, punishment-centered bureaucracy, and representative bureaucracy. Jaffe (2001) used these concepts in management: punishment centered management and representative bureaucracies. In the NPOs represented by participants in this study, communication about performance measurement moves from the caseworkers to the management to the funders. Feedback is primarily given to the caseworker only when there is a problem. Feedback from the funder to the agency suggests a punishment-centered environment. However, this was not the case within one agency in this study: staff members had positive perceptions of receiving adequate feedback and indicated that they were involved in the performance measurement process. Even though there were gaps in the

communication they received from supervisors, the staff members in this study were generally satisfied with the communication within the agency. This suggests a more representative type of management. What is clear is that there are multiple levels in which feedback is occurring, from the agency to the funder and from the agency to the staff members. But feedback is not done in a manner that supports organizational learning (Argyris, 1977; Senge, 2006).

In regard to performance measurement, one participant working as a caseworker stated that she did not get feedback from funders, but she did get feedback from her agency. She captures the predominant sentiment expressed by study participants throughout the interviews:

Myself, I don't really need that feedback because I know what I'm doing. I know what progress I have made with people. So, the numbers part of it, yes it's important and we have to have the numbers; we have to have it, the measures. But I also, for me, what drives me, is making that progress; empowering these people.

This statement again reiterates the commitment of the caseworkers in this study. But feedback and communication is more than about knowing when something is correct and when it is not.

Researchers have identified feedback as an essential component for organizational learning (Argyris, 1976; Senge, 2006). Organizational learning is the way in which organizations take experience and use that experience to react to the changing environment, improve their ability to provide services, and determine what practices are working. For organizational learning to occur, the information has to be complete; it has to involve all parties and it has to flow both up and down the organizational hierarchy (Argyris, 1977; Ashford & Cummings, 1983; Senge, 2006). This study suggests that there is inadequate communication and feedback for leaders within the human services field, both at the organizational level and at the funding level, to use

performance measurement for organizational learning. The organizations do not have the flexibility to change program guidelines to meet client needs, and the funding agencies do not have the information needed to assess what practices lead to better outcomes: at least not in a formal manner. It is beyond the scope of this study to understand how policy decisions are made at the funding level using performance measurement data, but this study does show that important information that the caseworkers have is not included in reports generated through performance measurement.

Limitations

The findings of this study that relate specifically to the way employees of a set of NPO agencies engaged in employment and training programs perceive performance measurement, I make no claim that these findings apply to other human service programs. There were four different employment and training (E&T) programs represented through this study, but the majority, 21 of 28 (75%), of the study participants worked in the Work Ready program. The conclusions and findings are limited in their representation of other employment and training programs.

This study focused on the perceptions of the service provider agency employees and there included no interviews with staff members from funding organizations. It was my intention to present views and experiences of the caseworkers and managers within the limits of their experiences. It was not my aim to present their perceptions in connection with outside sources. This study is not an attempt to criticize or indict any agencies or funding sources for the practices they use. Instead, I hope the results will inform interested parties in how caseworkers and managers of human services programs perceive and respond to performance measurement practices. Indeed, one study participant who had responsibilities in both management and direct

service to clients suggested that the state knows that the performance measurement information and what accounts for client outcomes “doesn’t really match up” but is constrained by federal requirements. Although this study is limited to participants’ perceptions and limited to their knowledge of what happens in their agencies, the results add to other literature that suggests that what matters in human service NPOs is not always what is captured through performance measurement.

Direction for Future Research

This study has provided findings on staff perceptions of performance measurement but future research is needed to more fully understand the affect that performance measurement has on organizations. Staff characteristics, training and communication are three issues that were repeatedly identified in this study as influencing the effect that performance measurement had on the staff member. These three issues should be looked at more in-depth and from a different perspective to improve our understanding of how performance measurement affects organizations and staff members.

Areas for future research need to include larger scale studies to look at the characteristics of the caseworkers that are being hired within the human services field. Are leaders looking to find staff members that are good at data entry or good at working with people? Are these two skill sets mutually exclusive? A topic for consideration is the shifting focus on data systems in the human services field and how that affects what type of people are attracted to the field, remain in the field, and their job satisfaction. If the trend toward more data collection within the human services field altering the work force itself, leaders need to be aware of the shifting needs and characteristics of their staff members within the context of their organizations. Shifts in staff characteristics may require different types of trainings, different division of labor, and other

considerations to effectively manage the work force to achieve an agency's mission. It is one achievement to have a perfect case file and 100% accuracy in data systems, it is a different achievement to have compassionate and caring caseworkers. Additional research can help leaders navigate the changing requirements of their workforce while still achieving their organization's mission.

Staff characteristics lead to broader gender issues within the human services field. It is not the intent of this study to analyze gender trends in the human services field; this study provides some suggestion that there is a link to the type of training that is provided and the perceptions of the type of workers within the field. In this study, I found that training was focused on data entry and system requirements over more interpersonal aspects required of caseworkers. The shift in trainings toward data collection as opposed to client interaction may be significant in several ways. It may show a gender bias toward a predominantly female work force. It could also prove to be undermining the human aspect of the goals of the programs by not recognizing the need for training on the social and psychological elements of effective programs. The sample size in this study is not large enough to make any conclusion, but as Table 3 shows, 78% of the study participants are women. In management, 71% of the staff members are women. The two study participants that had a direct role in data collection and training other staff members on the data collection were both women. Research on the perceptions of the characteristics of the work force and the types of trainings that are offered can help leaders understand their implicit biases and how to achieve more equitable work force expectations and create valid training requirements.

This study supports the value of the caseworkers themselves in influencing client outcomes, despite the program parameters and performance measurement limitations. However, when failure is not recognized or reported and when funders do not know what is happening

outside the program guidelines that positively affect the program's success, then funders will not know how to improve programs because they have incomplete information. The participants in this study related ways they exceeded program requirements to help clients. This may inflate the programs' effectiveness for clients over what would be found if programs were being administered strictly by the guidelines. Brown et al. (2015) found that NPO executives valued human capital as more important to organizational effectiveness than processes including performance measurement. Additional research needs to be done to understand how managers balance the performance measurement requirements affect on staff members and the competing pressure for reaching goals despite the requirements. Do the managers know the extent that caseworkers work outside the guidelines and turn a blind eye? Is there an informal structure in the organizations to accommodate competing needs? These questions reach into the complex theoretical issues of goal displacement, strain, and latent functions found in the work of Merton, Weber and others. This study focused on the perceptions of staff members but these perceptions can create larger issues at an organizational level that need to be understood.

The findings in this study suggest that training has become focused on correct data entry and coding. Future research with the purpose of understanding how training has changed over time and what agencies are doing independently of the program requirements would inform funders how training influences the overall achievement of the program goals. In this study, one of the reasons I chose to sample staff members of employment and training programs is because they do not have staff qualification requirements to which perceptions might be attributed. However, not having qualification requirements also means that it is up to agencies to find and train staff who can deliver the services effectively. The idea that funders and agencies are using training resources for data collection and entry rather than program delivery could have a strong

impact on the effectiveness of programs. Since this change has occurred over time, it could also be a factor in the perceived effectiveness of programs over time. Staff training should be a consideration in overall performance measurement program design, especially in programs that do not have specific staff qualifications. This research could also inform funders about the development of training to achieve consistent program delivery across agencies.

This study did not provide any answers on how much time staff members use for performance measurement. What I did find was it is not just the time spent on data collection and reporting that is consuming service delivery resources, but training resources are also being diverted to data collection. The focus of trainings has shifted from increasing caseworker's ability to provide quality service to training on proper data entry. In order to find an answer to how much time agencies spend on performance measurement a comprehensive time study would need to be completed. To fully understand the time cost of performance measurement, a study would need to make a distinction between performance measurement and data collection for other purposes such as client orientation and assessment. Such a study would also need to include the entire staff of the agency since support staff and, as I learned in this study, sometimes volunteers are used for data collection and reporting. A comprehensive time study could provide agencies and funders with information needed to develop best practices on efficient and meaningful data collection for reporting and performance measurement. It might also inform them on what practices lead to quality and consistency in data collection considering the range of uses includes assessment and eligibility.

I found the results on communication and feedback in this study intriguing. Individuals had conflicting perspectives regarding feedback that suggest that the desire and need for feedback depends on the context and the source of the feedback itself. There was an assumption by

participant caseworkers and managers that feedback from higher ups was given only when something was incorrect. As captured in one participant's assessment, "Basically, if you hear nothing, that's good." Future research should explore the existence of multiple feedback loops within an agency, the flow of positive and negative information in the loops, and actions taken by funders, agencies, and caseworkers that result from this feedback. According to Senge (2006), multiple feedback loops are required for an organization to be a learning organization. In NPOs, the constraints of funder requirements can impede that learning process. Understanding how feedback loops exist between agencies and clients and span across to the funders in NPOs could help inform agencies and funders on how to balance program improvement with compliance and accountability. It is also important to reveal to leaders the difference in how communication actually occurs versus their perception. Better communication within and between organizations and funders will allow performance measurement to be used as more than just a monitoring tool.

Another aspect of communication is the client. As I began this study, I was concerned that performance measurement would impede the ability of the caseworkers to interact positively with clients. This study demonstrated the vast amount of information that clients must divulge. Is there a bias in performance measurement data because clients who do not want to divulge personal information simply do not seek services? A study that looks at the reasons why people do not seek services or do not complete programs would provide better insight to the effect of this data collection on the client's willingness to receive help. In this study, one caseworker said that she had clients who simply did not come back to orientation after the first break. This may not be a widespread problem, but perhaps one worth exploring in the overall ability of organizations to reach out to the general population of low-income clients.

A common theme throughout my suggestions for future research is training. Training is being diverted from human issues to data needs. Types of training may be chosen because of gender biases. Trainings and qualifications of staff members vary between organizations. The variances of training, both required and received by case workers, has implications for the reliability of performance measurement as a tool for both improving programs and understanding a program's effectiveness. More research in all of the areas I have suggested will provide useful information in how human services can continue to provide a benefit to the clients they serve and sustain a strong work force within the human services field.

Implications for Policy and Practice

This study defines performance measurement as having three uses: monitoring, resource allocation, or improvement. There are delays that are caused by layers or reporting and the slow pace of policies shifts so the notion that performance measurement can be used for program improvement becomes suspect. For example, the 2014 CSBG report as cited in this dissertation was not published until 2018: 4 years after the work had been done. Since 2014 there have been major turning points in our culture on a national scale: the legalization of same sex marriage, the issues of how to record and recognize gender, and historically low unemployment rates. There are also changes at organizations: in this study fewer than half of the caseworkers had been with the agency for more than five years. There is no evidence in this study that performance measurement is used for resource allocation as shown by the responses to questions about funding. The responses of study participants were both vague and contradictory. That leaves monitoring. But using performance measurement for monitoring has shown to create issues of goal displacement and validity concerns.

This does not suggest that performance measurement does not have value. Accountability of agencies for the funding they receive is a real need. Determining what organizational practices are effective is important. What this study does show is that performance measurement has pushed the focus away from what is happening between the case worker and the client to what is happening between the caseworker and the computer. This study has found in the WR program that a client can have no barriers removed but the program can be considered a success because the paperwork is accurate and timely. This study has found that caseworkers are experiencing goal displacement demonstrated through their work arounds to agency policies and funding requirements. At the agency level, staff members are providing help to a client that they feel is against agency policy. At the funding level they are coding and entering data in a “creative way.”

So, what are policy makers and leaders to do? A core dilemma in organizations is that “we learn best from experience but we never directly experience the consequences of many of our most important decisions” (Senge, 2006, p.23). This is an apt description of the processes at work in this study. Often funders’ decisions are made in a sweeping manner that does not consider program differences (Pfeffer & Salancik, 1974). Do we really want our caseworkers to be spending 20% of their time doing data entry to provide someone somewhere with information when little actionable information ever flows back to the caseworker in a timely manner? There are better ways to use performance measurement that provides the accountability needed and also allows for organizations to learn and improve.

In this study, I included only agencies administering employment and training programs that had to meet the same performance measurement reporting standards. This minimized the variations across programs that occur in the human services field. I also chose to sample employees of Community Action Agencies (CAAs) that have the goal of helping people reach

self-sufficiency: employment and training programs are one way these agencies can accomplish this mission. CAAs are required to conduct performance measurement on all of the programs they administer through federally legislated reporting mechanisms. This combination of employment and training programs within CAAs provided the benefit of a study that considers the experiences and perceptions of employees using two separate performance measurement systems within one agency. This study did not intend to imply any bias toward CAA administering human services programs. However, the findings indicate that staff perceive performance measurements done through CAA funding as providing more meaningful information on client outcomes than the reporting requirements for the employment and training programs CAAs administer. This also demonstrates that at least one performance measurement system exists which can collect meaningful outcomes.

In monitoring for data collection, determining if the service provider is entering data in correctly, have they completed reports on time and is the coding right and consistent, we are missing the human element. While the data collection process is being scrutinized the actual quality of service delivery is overlooked. As the adage states “what gets measured gets done” and “what gets monitored gets done correctly” (Saylor, 1994). Is the concern for correct data entry driving how we hire and train caseworkers? Organizations have had two types of accountability: in the 1960’s accountability centered on fiscal elements of the organization, (Martin & Frahm, 2010) current emphasis on accountability has focused on performance measurement. Perhaps it is time to have three types of accountability: fiscal, performance, and systems. Fiscal monitoring entails determining if the funds themselves are being used legally and as intended. Through audits or program specific monitoring, these types of checks on an organization are routine. Current practices in performance measurement appear to address how guidelines are being followed and if

data is recorded correctly. Perhaps an assumption is being made that this type of reporting is actually an indicator that outcomes are being achieved: conflating correct data with correct action. This study shows that this is not always the case. The data collection being handled correctly and clients being cared for correctly are two different work tasks. Monitoring should also be looking at these differently. Performance measurement should whether the data systems are being used correctly as well as the effective engagement of the clients.

I am not suggesting that programs have to meet program fidelity standards in performance measurement, just the opposite. Program evaluation requires a level of program fidelity: the adherence to the original intent of the program design when the program is delivered to clients. (Mowbray, Bybee, Holter & Lewandowski, 2006). This requires programs to be administered exactly as it is designed. In this study, I saw a wide variety of ways that agencies implement programs. Why not use that as a strength? Agencies already gather extensive information on the clients; we need information on the agencies, the staff, caseload, and the local environment. What factors contribute to one agency's success while another agency struggles? Understanding how those differences affect the client outcomes could allow funders and human service professionals to determine best practices. Funding agencies need to separate the process of monitoring agencies for accountability compliance from the process of determining what makes programs effective.

This is possible and not a new idea. Carver and Carver (2009) call this type of management the *Policy Governance Model*. The phrase they use is "freedom within boundaries" (Carver & Carver, 2009, p. 13). By providing broad limits and allowing the agencies to function within those limits, management creates an environment that allows day-to-day decisions to be made locally. Compliance is limited to not going outside specific boundaries. This type of accountability might allow funders to get information that is more complete on client outcomes.

Performance measurement then becomes a tool to not only evaluate effectiveness but to also determine why a program is effective. It would also have the benefit of freeing policy makers from developing extensive program guidelines that caseworkers then work around. Give an agency free reign on how they can work with a client, and see what they do. Give broad parameters, and see what happens. Then provide best practices, choices, and support. That is the only way program improvement can occur in any meaningful way. In this study, there is an irony in the wealth of client data from the orientation and assessment processes that funders have access to and the performance measurement data that is collected. Agencies have an abundance of baseline data about clients, but no indication from staff members that funders are using this information to determine the change that occurs for clients through participation in the program.

Policy makers and leaders at all levels should come to the realization of what performance measurement can and cannot do in its current state. In this study I could not find any evidence that performance measurement is used in any substantive way to guide resource allocation. Its use for monitoring is heavily focused on correct data entry and coding. The use of performance measurement for program improvement is illusive because feedback to caseworkers does not include information on how to improve service delivery. What this study does find is that caseworkers are generally dedicated to helping clients and there is a wealth of data that exists about those clients. Organizational leaders can use these strengths and resources to improve programs but it will take a shift in thinking. Performance measurement should be viewed as more than just a monitoring tool and data repository. There is talk of program improvement and even a hope that it occurs, but leaders need to take responsibility to see that it does happen and that those improvements are communicated to others in the field. Not just feel good stories, but functional changes that can improve everyone's efforts to help those in need.

Conclusions

From the passage of the Government Performance and Results Act (1993) by President Clinton to the Government Performance and Results Modernization Act (2010) signed by President Obama, the federal government has mandated the use of performance measurement. As I complete this dissertation, the White House has just released their 2019 fiscal year budget proposal. It shows that performance measurement continues to be relevant and important. The concepts are even in the title: *An American Budget; Efficient, Effective, Accountable* (OMB, 2018). In the introductory section, “Modernizing Government for the 21st Century” (p. 7), this budget document highlights the need for quality performance measurement: “The public lacks sufficient opportunities to give feedback on Federal programs and services, making it harder to identify weaknesses and make improvements. Smarter use of data and evidence is needed to orient decisions and accountability around service and results” (OMB, 2018, p. 7). As this study shows, communication between funders, agencies, and clients is a key element for performance measurement and refining delivery of services. That communication needs to be open and honest to allow for real program improvement. If what one caseworker claims is true, that “they have no idea what we do,” performance measurement cannot exist in any meaningful way.

The current White House budget proposal also claims, “Service delivery sometimes suffers due to checking unnecessary bureaucratic boxes. Managers need greater discretion to execute programs effectively, foster the highest-performing workforce, and solve real-time problems” (OMB, 2018, p. 7). This discretion and flexibility needs to filter down to the agencies administering the programs. When funders use performance measurement for compliance and monitoring, focusing on the right code over the best course of action, we lose the ability to change and to learn from what is happening. The successes of human services, by their very name, are

dependent on the actions of individuals not the existence of guidelines, and on relationships, not data. Caseworkers should not be afraid to say they helped a client; successes should be celebrated.

This year's White House (2019) budget proposal continues, "tax payer dollars must go to effective programs that efficiently produce results" (p. 8). This study demonstrates that performance measurement has the potential to help nonprofit and government organizations be more effective without taking away from their mission. Government and nonprofit organizations have created complicated software systems, databases, forms, and procedures all in an effort to collect and report large amounts of information about program participants. That is only the first step; now it is time to take the next step and use that information to make programs more effective and thereby improving organizational effectiveness.

References

- Abramovitz, M. (1992). The Reagan legacy: Undoing class, race and gender accords. *The Journal of Sociology and Social Welfare*, 19(1), 91-110.
- Acker, J. (1990). Hierarchies, jobs, bodies: A theory of gendered organizations. *Gender & Society*, 4(2), 139-158.
- Acker, J. (2006). Inequality regimes: Gender, class, and race in organizations. *Gender & Society*, 20(4), 441-464.
- Agnew, R. (2001) Building the foundation of general strain theory: Specifying the types of strain most likely to lead to crime and delinquency. *Journal of Research in Crime and Delinquency*, 38(4), 319-361.
- Alexander, C., & Charles, G. (2009). Caring, mutuality and reciprocity in social worker-client relationships rethinking principles of practice. *Journal of Social Work*, 9(1), 5-22. doi: 11.1177/1468017308098420
- Amabile, T. M. (1979). Effects of external evaluation on artistic creativity. *Journal of Personality and Social Psychology*, 37(2), 221.
- Appelrouth, S., & Edles, L. D. (Eds.). (2011). *Sociological theory in the contemporary era: Text and readings* (2nd ed.). Thousand Oaks, CA: Pine Forge Press.
- Arenson, K. W. (1995, April 4). Former United Way chief guilty in theft of more than \$600,000. *New York Times*, A1.
- Argote, L., & Miron-Spektor, E. (2011). Organizational learning: From experience to knowledge. *Organization Science*, 22(5), 1123-1137.
- Argyris, C. (1976). Single-loop and double-loop models in research decision making. *Administrative Science Quarterly*, 21(3), 363-375.

- Argyris, C. (1977). Double loop learning in organizations. *Harvard Business Review*, 55(5), 115-125.
- Ashford, S. J., & Cummings, L. L. (1983). Feedback as an individual resource: Personal strategies of creating information. *Organizational Behavior and Human Performance*, 32(3), 370-398.
- Ashforth, B. E., & Humphrey, R. H. (1995). Emotion in the workplace: A reappraisal. *Human Relations*, 48(2), 97-125.
- Bazeley, P. (2009). Analysing qualitative data: More than 'identifying themes'. *Malaysian Journal of Qualitative Research*, 2(2), 6-22.
- Behn, R. D. (2003). Why measure performance? Different purposes require different measures. *Public Administration Review*, 63(5), 586-606.
- Benjamin, L. M. (2008). Bearing more risk for results performance accountability and nonprofit relational work. *Administration & Society*, 39(8), 959-983.
- Benjamin, L. M. (2012). Nonprofit organizations and outcome measurement from tracking program activities to focusing on frontline work. *American Journal of Evaluation*, 33(3), 431-447. doi:10.1177/1098214012440496
- Benjamin, L. M., & Campbell, D. C. (2015). Nonprofit performance accounting for the agency of clients. *Nonprofit and Voluntary Sector Quarterly*, 44(5), 988-1006.
doi:10.1177/0899764014551987
- Blalock, A. B. (1999). Evaluation research and the performance management movement from estrangement to useful integration? *Evaluation*, 5(2), 117-149.
- Bloomberg, L. D., & Volpe, M. (2008). *Completing qualitative research: A roadmap from beginning to end*. Thousand Oaks, CA: Sage Publications.

- Bradley, E. H., Curry, L. A., & Devers, K. J. (2007). Qualitative data analysis for health services research: developing taxonomy, themes, and theory. *Health Services Research, 42*(4), 1758-1772.
- Broidy, L., & Agnew, R. (1997). Gender and crime: A general strain theory perspective. *Journal of Research in Crime and Delinquency, 34*(3), 275-306.
- Britton, D. M. (2000). The epistemology of the gendered organization. *Gender & Society, 14*(3), 418-434.
- Bromley, P., & Meyer, J. W. (2017). "They are all organizations": The cultural roots of blurring between the nonprofit, business, and government sectors. *Administration & Society, 49*(7), 939-966.
- Brown, W. A., Andersson, F. O., & Jo, S. (2015). Dimensions of capacity in nonprofit human service organizations. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations, 1-24*. doi:10.1007/s11233-01596338
- Carlson, J., Kelley, A. S., & Smith, K. (2010). Government performance reforms and nonprofit human services: 20 years in Oregon. *Nonprofit and Voluntary Sector Quarterly, 39*(4), 630-652. doi:10.1177/0899764009360824
- Carmen, J. G. & Fredricks, K. A. (2010). Evaluation capacity and nonprofit organizations: Is the glass half-empty or half-full? *American Journal of Evaluation 31*(1), 84-104.
doi:10.1177/1098214009352361
- Carnochan, S., Samples, M., Myers, M., & Austin, M. J. (2014). Performance measurement challenges in nonprofit human service organizations. *Nonprofit and Voluntary Sector Quarterly, 43*(6), 1014-1032. doi:10.1177/0899764013508009

- Carver, J., & Carver, M. M. (2009). *A Carver Policy Governance Guide, The Policy Governance Model and the Role of the Board Member* (Vol. 1). San Francisco, CA: John Wiley & Sons.
- Christensen, R. A., & Ebrahim, A. (2006). How does accountability affect mission? The case of a nonprofit serving immigrants and refugees. *Nonprofit Management and Leadership, 17*(2), 195-209.
- Corbin, J. J. (1999). A study of factors influencing the growth of nonprofits in social services. *Nonprofit and Voluntary Sector Quarterly, 28*(3), 296-314.
- Courtney, M. E., Needell, B., & Wulczyn, F. (2004). Unintended consequences of the push for accountability: The case of national child welfare performance standards. *Children and Youth Services Review, 26*(12), 1141-1154. doi:10.1016/j.childyouth.2004.05.005
- Creswell, J.W. (2009). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage Publications.
- Crittenden, W. F. (2000). Spinning straw into gold: The tenuous strategy, funding, and financial performance linkage. *Nonprofit and Voluntary Sector Quarterly, 29*(suppl 1), 164-182.
- Danziger, S., Sandefur, G. D., & Weinberg, D. H. (Eds.). (1994). *Confronting poverty: Prescriptions for change*. Harvard University Press.
- DeHoog, R. H. (1985). Human Services Contracting Environmental, Behavioral, and Organizational Conditions. *Administration & Society, 16*(4), 427-454.
doi:10.1177/009539978501600403
- De Tocqueville, A. (1935/2007). *Democracy in America*. In Calhoun, Craig et al. (Eds.). *Classical Sociological Theory* (55-71). San Francisco: Blackwell

- Dicke, L. A. (2002). Ensuring Accountability in Human Services Contracting Can Stewardship Theory Fill the Bill? *The American Review of Public Administration*, 32(4), 455-470.
- Dicke, L. A., & Ott, J. S. (1999). Public agency accountability in human services contracting. *Public Productivity & Management Review*, 22(4), 502-516.
- DiMaggio, P. J., & Powell, W.W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147-160.
- Drucker, P. F. (2011). *The five most important questions you will ever ask about your organization* (Vol. 90). San Francisco, CA: John Wiley & Sons.
- Dubnick, M. (2003). Accountability and the promise of performance: In search of the mechanisms. *Public Performance & Management Review*, 28(3), 376-417.
- Ebrahim, A.S. (2003). Making sense of accountability: Conceptual perspectives for northern and southern nonprofits. *Nonprofit Management and Leadership*, 14(2), 191-212.
- Ebrahim, A.S. & Rangan, V. K. (2010). The Limits of Nonprofit Impact: A Contingency Framework for Measuring Social Performance. Harvard Business School General Management Unit Working Paper No. 10-099; Harvard Business School Marketing Unit Working Paper No. 10-099. Available at SSRN: <https://ssrn.com/abstract=1611810> or <http://dx.doi.org/10.2139/ssrn.1611810>
- Farnworth, M., & Leiber, M. J. (1989). Strain theory revisited: Economic goals, educational means, and delinquency. *American Sociological Review*, 54(2), 263-274.
- Froelich, K. A. (1999). Diversification of revenue strategies: Evolving resource dependence in nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly*, 28(3), 246-268.
doi:10.1177/0899764099283002

feedback. Business Dictionary.com. Retrieved September 10, 2017, from Business Dictionary .com website: <http://www.businessdictionary.com/definition/feedback.html>

for profit organization. BusinessDictionary.com. Retrieved June 10, 2018, from Businessdictionary.com website: [//www.businessdictionary.com/definition/for-profit-organization.html](http://www.businessdictionary.com/definition/for-profit-organization.html).

Glisson, C. A. & Martin, P. Y. (1980). Productivity and efficiency in human service organizations as related to structure, size and age. *Academy of Management Journal* 23(1), 21-37.

Gouldner, A. W. (1954). *Patterns of industrial bureaucracy*. New York: The Free Press

Government Performance and Results Act of 1993, S. 20, 103rd Cong (1993)

GPRA Modernization Act of 2010, H.R. 2142, 111th Cong. (2011)

Grandey, A. A. (2000). Emotional regulation in the workplace: A new way to conceptualize emotional labor. *Journal of Occupational Health Psychology*, 5(1), 95-110.

Grobman, G. M. (2008). *The Pennsylvania nonprofit handbook* (8th ed.). Harrisburg, PA: White Hat Communications.

Haines, E. L., Deaux, K., & Lofaro, N. (2016). The times they are a-changing... or are they not? A comparison of gender stereotypes, 1983–2014. *Psychology of Women Quarterly*, 40(3), 353-363.

Halachmi, A. (2005). Performance measurement: test the water before you dive in. *International Review of Administrative Sciences*, 71(2), 255-266.

Hall, P. D. (2003). *A solution is a product in search of a problem: A History of Foundations and Evaluation Research*. Hauser Center for Nonprofit Organizations, Kennedy School of Government; Harvard University.

- Hall, P. D. (2006). A historical overview of philanthropy, voluntary associations, and nonprofit organizations in the United States, 1600-2000. *The Nonprofit Sector: A Research Handbook*, (2), 32-65.
- Hasenfeld, Y. (2010). The attributes of human service organizations. *Human Services as Complex Organizations*, (2), 9-32.
- Hatry, H. P. (2002). Performance measurement: Fashions and fallacies. *Public Performance & Management Review*, 25(4), 352-358.
- Hatry, H. P. (2013). Sorting the relationships among performance measurement, program evaluation, and performance management. *New Directions for Evaluation*, 2013(137), 19-32. doi:10.1002/ev.20043
- Hendricks, M., Plantz, M. C., & Pritchard, K. J. (2008). Measuring outcomes of United Way–funded programs: Expectations and reality. *New Directions for Evaluation*, 2008(119), 13-35. doi:10.1002/ev.266
- Herman, R. D., & Renz, D. O. (1997). Multiple constituencies and the social construction of nonprofit organization effectiveness. *Nonprofit and Voluntary Sector Quarterly*, 26(2), 185-206.
- Hochschild, A. R. (2012). *The managed heart: Commercialization of human feeling*. University of California Press.
- Hofer, R. (2000). Accountability in action?: Program evaluation in nonprofit human service agencies. *Nonprofit Management and Leadership*, 11(2), 167-177.
- Hopkins, B.R., (2016), Wounded Warrior Project Takes PR Hit. *Nonprofit Counsel*, 33(7). doi: 10.1002/npc.30198

- Jacobs, R., & Washington, C. (2003). Employee development and organizational performance: a review of literature and directions for future research. *Human Resource Development International*, 6(3), 343-354. doi:10.1080/13678860110096211
- Jaffee, D. (2001). *Organization theory: tension and change*. New York: McGraw-Hill Higher Education.
- Kaplan, R. S. (2001). Strategic performance measurement and management in nonprofit organizations. *Nonprofit Management and Leadership*, 11(3), 353-370.
- Kazbour, R. R., McGee, H. M., Mooney, T., Masica, L., & Brinkerhoff, R. O. (2013). Evaluating the impact of a performance-based methodology on transfer of training. *Performance Improvement Quarterly*, 26(1), 5-33.
- Kim, S. E. (2005). Balancing competing accountability requirements: Challenges in performance improvement of the nonprofit human services agency. *Public Performance & Management Review*, 29(2), 145-163.
- Kramer, R. M. (1994). Voluntary agencies and the contract culture: " Dream or nightmare?" *The Social Service Review*, 68(1),33-60.
- Kruml, S. M., & Geddes, D. (2000). Exploring the dimensions of emotional labor: The heart of Hochschild's work. *Management Communication Quarterly*, 14(1), 8-49.
- Kuhn, M. (1971). Bureaucratic man: A portrait and an interpretation. *American Sociological Review* 36(3), 461-474.
- LeRoux, K., & Wright, N. S. (2010). Does performance measurement improve strategic decision making? Findings from a national survey of nonprofit social service agencies. *Nonprofit and Voluntary Sector Quarterly* 39(4),571-587. doi:10.1177/0899764009359942

- Levitt, B., & March, J. G. (1988). Organizational learning. *Annual Review of Sociology*, 14(1), 319-338.
- Liket, K. C., Rey-Garcia, M., & Maas, K. E. (2014). Why aren't evaluations working and what to do about it a framework for negotiating meaningful evaluation in nonprofits. *American Journal of Evaluation*, 35(2), 171-188. doi:10.1177/1098214013517736
- MacIndoe, H., & Barman, E. (2012). How organizational stakeholders shape performance measurement in nonprofits: Exploring a multidimensional measure. *Nonprofit and Voluntary Sector Quarterly*, 42(4) 716-738. doi;10.1177/0899764012444351
- Martin, L. L., & Frahm, K. (2010). Changing nature of accountability in administrative practice. *The Journal of Sociology and Social Welfare*, 37(1), 137-148.
- Martin, L. L., & Kettner, P. M. (1997). Performance measurement: The new accountability. *Administration in Social Work*, 21(1), 17-29.
- Maxwell J. A., (2005). *Qualitative research Design: An interactive approach* (2nd Ed.). Thousand Oaks CA: Sage Publications
- McElroy, M. W. (2000). Integrating complexity theory, knowledge management and organizational learning. *Journal of Knowledge Management*, 4(3), 195-203.
- McHargue, S. K. (2003). Learning for performance in nonprofit organizations. *Advances in Developing Human Resources*, 5(2), 196-204.
- McKeever, B. S., & Pettijohn, S. L. (2014). *The nonprofit sector in brief 2014*. Washington, DC: Urban Institute.
- Melkers, J. (2006). On the road to improved performance: Changing organizational communication through performance management. *Public Performance & Management Review*, 30(1), 73-95.

- Merton, R. K. (1936). The unanticipated consequences of purposive social action. *American Sociological Association*, 1(6), 894-904.
- Merton, R. K. (1938). Social structure and anomie. *American Sociological Review*, 3(5), 672-682.
- Merton, R. K. (1957/2007). Manifest and latent functions. In Calhoun, Craig et al. (Eds.). *Classical Sociological Theory* (441-448). San Francisco, CA: Blackwell
- Merton, R. K. (1957/2011). Bureaucratic structure and personality. In J.M. Shafritz, J.S.Ott & Y.S. Jang, (Eds.), *Classics of Organizational Theory* (pp. 107-115). Boston, MA: Wadsworth.
- Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340-363.
- Mitchell, G. E. (2012). The construct of organizational effectiveness: Perspectives from leaders of international nonprofits in the United States. *Nonprofit and Voluntary Sector Quarterly*, 42(2), 324-345. doi:10.1177/0899764011434589
- Monette, D. R., Sullivan, T. J., & DeJong, C. R. (2011). *Applied social research: A tool for the human services* (8th ed.). Belmont, CA: Cengage Learning.
- Mowbray, C. T., Bybee, D., Holter, M., & Lewandowski, L. (2006). Validation of a fidelity rating instrument for consumer-operated services. *American Journal of Evaluation*, 27(1), 9-27.
- Moynihan, D. P. (2006). What do we talk about when we talk about performance? Dialogue theory and performance budgeting. *Journal of Public Administration Research and Theory*, 16(2), 151-168. doi:10.1093/jopart/muj003
- Moynihan, D. P. (2006a). Managing for results in state government: Evaluating a decade of reform. *Public Administration Review*, 66(1), 77-89.

- Moynihan, D. P., Fernandez, S., Kim, S., LeRoux, K. M., Piotrowski, S. J., Wright, B. E., & Yang, K. (2011). Performance regimes amidst governance complexity. *Journal of Public Administration Research and Theory*, 21(suppl 1), i141-i155.
- Murphy, D. M. (2007). *Beyond accountability: An empirical study of the factors associated with the use of evaluation for organizational learning in North Carolina's nonprofit sector*. (Doctoral Dissertation, North Carolina State University).
Retrieved from <https://www.lib.ncsu.edu/resolver/1840.16/4298>.
- National Community Action Partnership, (2018). *About Community Action Partnership*.
Retrieved from: <https://communityactionpartnership.com/menus/about-us.html>.
- Organ, D. W. (1988). A restatement of the satisfaction-performance hypothesis. *Journal of Management*, 14(4), 547-557.
- O'Neill, O. (2002). *A question of trust: The BBC Reith Lectures 2002*. Cambridge University Press.
- Ott, J. S., Shafritz, J. M., & Jang, Y. S. (2011). *Classic readings in organization theory* (7th Ed.). Boston: Wadsworth/Cengage Learning.
- Packard, T. (2009). Staff perceptions of variables affecting performance in human service organizations. *Nonprofit and Voluntary Sector Quarterly*, 39(6),971-990.
doi:11.1177/0899764009342896
- Pandey, S. K., Coursey, D. H. & Moynihan, D. P. (2006). Organizational effectiveness and bureaucratic red tape: a multimethod study. *Public Performance Management and Management Review*, 30(3), 398-425. doi:10.2753/PMR1530-9576300305

- Pennsylvania Department of Community and Economic Development, Center for Community Services. (2014). *Reporting Requirements Directive C2014-05*. Retrieved from http://www.thecaap.org/file_download/0aca4437-7f2c-4688-be56-c486a0a9fa40.
- Pennsylvania Department of Human Services (2018) Employment programs for people who receive benefits. Retrieved from <http://www.dhs.pa.gov/publications/infographics/employmentservices/index.htm>.
- Pfeffer, J., & Salancik, G. R. (1974). Organizational decision making as a political process: The case of a university budget. *Administrative Science Quarterly*, 19(2)135-151.
- Powers, C.H. (2010). *Making sense of social theory: A practical introduction*. (2nd ed.). Lanham MD: Rowman & Littlefield Publishers.
- Powell, W. W., & Colyvas, J. A. (2008). Microfoundations of institutional theory. In R. Greenwood, C. Oliver, K. Sahin & R. Suddaby (Eds.), *The Sage Handbook of Organizational Institutionalism*, (pp. 276-298). Thousand Oaks CA: Sage Publications.
- Quinn, R. E., & Rohrbaugh, J. (1983). A spatial model of effectiveness criteria: Towards a competing values approach to organizational analysis. *Management Science*, 29(3), 363-377.
- Russ-Eft, D., & Preskill, H. (2009). *Evaluation in organizations a systematic approach: A systematic approach to enhancing learning, performance, and change*. New York, NY: Basic Books.
- Sanger, M. B. (2005). Successful strategies for engaging citizens in performance measurement. *National Civic Review*, 94(3), 53-54.
- Sanger, M. B. (2008). From measurement to management: Breaking through the barriers to state and local performance. *Public Administration Review*, 68(s1), S70-S85.

- Sanger, M.B., 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(4), 711-728.
- Saylor, W. G. (1994). The design, development and maintenance of strategic support systems for correctional agencies. *International Symposium on Criminal Justice Information Systems and Technology*, Washington DC.
- Scott, B. A., & Barnes, C. M. (2011). A multilevel field investigation of emotional labor, affect, work withdrawal, and gender. *Academy of Management Journal*, 54(1), 116-136.
- Selznick, P. (1996). Institutionalism" old" and" new". *Administrative Science Quarterly*, 42(2), 270-277.
- Senge, P.M. (2006). *The fifth discipline: The art and practice of the learning organization*. New York, NY: Doubleday.
- Steers, R. M. (1975). Problems in the measurement of organizational effectiveness. *Administrative Science Quarterly*, 20(2), 546-558.
- Suddaby, R. (2010). Challenges for institutional theory. *Journal of Management Inquiry*, 19(1), 14-20. doi:10.1177/1056492609347564
- Temple, Harrisburg, 2018. *Credential for strengths-based family workers*. Retrieved from [www.https://www.harrisburg.temple.edu/continuing-education/credential-strengths-based-family-workers](https://www.harrisburg.temple.edu/continuing-education/credential-strengths-based-family-workers).
- Tetlock, P. E. (1985). Accountability: The neglected social context of judgment and choice. *Research in Organizational Behavior*, 7(1), 297-332.

- Thomson, D. E. (2010). Exploring the role of funders' performance reporting mandates in nonprofit performance measurement. *Nonprofit and Voluntary Sector Quarterly*, 39(4), 611-629.
- Tompkins, J. (2005). *Organization theory and public management*. Belmont, CA: Thomson Wadsworth.
- Udy Jr, S. H. (1959). "Bureaucracy" and "Rationality" in Weber's organization theory: An empirical study. *American Sociological Review* 24(6), 791-795
- United States Government Accountability Office, (May 2011) *Performance Measurement and Evaluation; Definitions and relationships*. GAO -11-646SP. Retrieved from <https://www.gao.gov/special.pubs/gg98026.pdf>.
- United States Census Bureau, *2010 – 2014 American Community Survey 5-year Estimates* [Data file]. Retrieved from https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_14_5YR_DP03&src=pt.
- United States Department of the Treasury, Bureau of the Fiscal Services. (2016). *The Debt to the Penny and Who Holds It*. Retrieved from <http://www.treasurydirect.gov/NP/debt/current>.
- United States General Accounting Office, *The Government Performance and Results Act: 1997 Governmentwide Implementation Will Be Uneven, June 1997*, GAO/GGD-97-109.
- United States Department of Health and Human Services, Administration for Children and Families, Office of Community Services (2015) OCS Im 138: State Establishment of Organizational Standards fir CSBG Eligible Entities. Retrieved from <http://nascsp.org/Performance-Management/967/Organizational-Standards.aspx?iHt=33>

- United States Department of Health and Human Services, Administration for Children and Families, Office of Community Services United States Office of Community Services (2018). *CSBG Report to Congress FY 2014*. Retrieved from <https://www.acf.hhs.gov/ocs/resource/csbg-report-to-congress-fy-2014>.
- United States Office of Management and Budget, (2018) *An American Budget: Efficient, Effective, Accountable Fiscal year 2019*. Retrieved from <https://www.whitehouse.gov/omb/budget/>
- Van Slyke, D. M. (2003). The mythology of privatization in contracting for social services. *Public Administration Review* (63)3, 296-315.
- Weber, M. (1914). The types of legitimate domination. In Calhoun, C., Gerteis, J., Moody, J., Pfaff, S., & Virk, I. (Eds.) (2007). *Classical sociological theory* (2nd ed.). (pp.256 - 263) Malden: Blackwell.
- Welsh, E. (2002). Dealing with Data: Using NVivo in the Qualitative Data Analysis Process [12 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 3(2), Art. 26, Retrieved from <http://www.qualitative-research.net/index.php/fqs/article/view/865/1880>.
- Yukl, G. (1999). An evaluative essay on current conceptions of effective leadership. *European Journal of Work and Organizational Psychology*, 8(1), 33-48.

Appendix A

Interview Protocol

SPACES REMOVED FROM ORIGINAL DOCUMENT

Date

Time

ORG

Hi, my name is Stacey Tice and I'm a human services worker and a doctoral student at Indiana University of PA. Your organization Executive Director has allowed me to ask for your help with my research I would like to talk with you, as someone who work with employment and training programs, in a one to one interview for just an hour for my dissertation study. I'm interested in your experiences with collecting the performance measure information we gather, and your views about it , The interviews would be confidential and your information would not be reported to anyone in a way that could identify you, It's mu hope that we can learn more about ow the performer measurement process actually happened so we can inform funders and policy makers about ways to collect and use this information effectively. If you think you might be willing to help me, please respond an I'll get in touch with you to provide more information and to set up an interview.

Interview Script

1. So as part of your job, I understand that you collect information about your employment and training clients and programs. What's you experience with that?
 - a. Tell me about the types of information and how you collect it?
 - b. Do you run into any difficulties gathering the necessary data?
2. How does collecting this information affect how you do your job?
 - a. Tell me about the time it takes?
 - b. How does that impact the help you can give clients?
3. What's your sense of how this information gets used in terms of programs for service delivery to clients?
 - a. What types of reports are generated from the information?
 - b. How is funding affected?
4. Tell me about the training you get for collecting and using this information?
 - a. What types of trainings do you wish you received?
5. Tell me about how you are involved in the process, shat input do you have?
 - a. Do you track more information than is required?
6. From your perspective, what would it take to make the information you collect more meaningful?
7. What type of feedback do you get from the information you collect?
 - a. What type of feedback would you like to get?

Demographics:

Gender: Man: _____ woman: _____ Other: _____

Race/ethnicity: _____

What is your highest level of education?

_____ Less than HS

_____ High school or GED

_____ Some College

_____ 4 Year College degree

_____ Graduate Degree

Other _____

How long have you been working for this agency? _____

How long have you been working in the human services field? _____

How long have you been working in employment and training? _____

How many people work in your Employment and Training programs? _____

Appendix B
Informed Consent Form



Indiana University of Pennsylvania
www.iup.edu

Institutional Review Board for the
Protection of Human Subjects
School of Graduate Studies and Research
Stright Hall, Room 113
210 South Tenth Street
Indiana, Pennsylvania 15705-1048

P 724-357-7730
F 724-357-2715
irb-research@iup.edu
www.iup.edu/irb

October 31, 2016

Stacey Tice
503 South Juliana Street
Bedford, PA 15522

Dear Ms. Tice:

Your proposed research project, "Staff Perceptions of Impact of Performance Measurement Requirements on Program Delivery and Organizational Effectiveness in Nonprofit Human Service Agencies," (Log No. 16-243) has been reviewed by the IRB and is approved. In accordance with 45CFR46.101 and IUP Policy, your project is exempt from continuing review. This approval does not supersede or obviate compliance with any other University requirements, including, but not limited to, enrollment, degree completion deadlines, topic approval, and conduct of university-affiliated activities.

You should read all of this letter, as it contains important information about conducting your study.

Now that your project has been approved by the IRB, there are elements of the Federal Regulations to which you must attend. IUP adheres to these regulations strictly:

1. You must conduct your study exactly as it was approved by the IRB.
2. Any additions or changes in procedures must be approved by the IRB before they are implemented.
3. You must notify the IRB promptly of any events that affect the safety or well-being of subjects.
4. You must notify the IRB promptly of any modifications of your study or other responses that are necessitated by any events reported in items 2 or 3.

The IRB may review or audit your project at random *or* for cause. In accordance with IUP Policy and Federal Regulation (45CFR46.113), the Board may suspend or terminate your project if your project has not been conducted as approved or if other difficulties are detected

Although your human subjects review process is complete, the School of Graduate Studies and Research requires submission and approval of a Research



Indiana University of Pennsylvania

Administration & Leadership Studies, Nonprofit & Public Sectors, Ph.D. Program
Department of Sociology 724-357-2956
102 McElhane Hall Fax: 724-357-4842
Indiana, Pennsylvania 15705-1087 www.iup.edu/sociology/als

STUDY OF STAFF PERCEPTIONS OF PERFORMANCE MEASURES IN NONPROFIT HUMAN SERVICE AGENCIES

INFORMED CONSENT FORM

You are invited to participate in this research study. The following information is provided in order to help you to make an informed decision whether or not to participate. If you have any questions please do not hesitate to ask. You are eligible to participate because you are an employee of a Community Action Agency and are involved with employment and training programs.

The purpose of this study is to learn more about how staff members in nonprofit human service organizations experience and view the processes of collecting and reporting information required for accountability measures. It is hoped that the information gained from this study may help us to inform funders' and policymakers' efforts to make performance measures an effective tool in service delivery and organizational functioning.

Your participation is confidential. Participation in this study involves a confidential individual interview of about an hour. The information you provide will be combined with the responses of other participants from across Pennsylvania for analysis. Information from the study will not be reported in a way that anyone could identify you or your agency/program. No information from you or other individuals at your agency will be reported back to the agency

For accuracy in transcribing our interview, I would like your permission to audio record it. Once the interview is transcribed, I will destroy the audio recording. If you grant permission for me to record our interview, please initial here: _____

Once data analyses are complete and my dissertation report is approved, I will destroy any information I have that would link you with your responses in this study, so that the data is completely anonymous and no one will be able to identify you as a participant in the study or with your responses. I want you to know that your confidentiality, and that of your agency, is important and I am making careful efforts to protect them.

Your participation is voluntary. You are free to decide not to participate in this study or to withdraw at any time. I will not tell anyone at your agency whether you did or did not participate in the study. If you choose to participate, you may withdraw at any time by notifying me, Stacey Tice. Upon your request to withdraw from the study, all information pertaining to you will be destroyed. If you choose to participate, all information will be held in strict confidence and will have no bearing on your employment or agency. Your responses will be considered only in combination with those from other participants and analyzed for patterns. The results of this research study may be published in scientific journals or presented at scientific meetings, but your identity will not be revealed and your agency will not be identifiable in reports so they remain strictly confidential.

Page 1 of 3

Statement of Risks

There are no known risks with participating in this study.

Statement of Benefits

There are no direct benefits to participating in this study. You, as a participants, may appreciate a chance to express your views about an accountability reporting system in which you may typically have little to say. It may be gratifying to have your voice heard and to contribute to a study that aims to inform policy and practice.

If you are willing to participate in this study, please sign the statement on the following page and return it to Stacey Tice (either in person, by email, or via postal mail to the address below). Please keep an unsigned copy for your records.

Thank you for considering this request to participate in this study.

Sincerely,

Stacey Tice, Ph.D. Candidate, Principal Investigator
Administration & Leadership Studies, Nonprofit and Public Sectors Ph.D. Program
Department of Sociology
Indiana University of Pennsylvania
McElhaney Hall, Room 102; 441 North Walk; Indiana, PA 15705
s.l.tice@iup.edu

J. Beth Mabry, Ph.D, Associate Professor, Faculty Sponsor
Department of Sociology
Indiana University of Pennsylvania
McElhaney Hall, Room 102; 441 North Walk; Indiana, PA 15705
mabry@iup.edu

*This project has been approved by the Indiana University of Pennsylvania
Institutional Review Board for the Protection of Human Subjects (Phone: 724/357-7730).*



Indiana University of Pennsylvania

Administration & Leadership Studies, Nonprofit & Public Sectors, Ph.D. Program

Department of Sociology

102 McElhaney Hall

Indiana, Pennsylvania 15705-1087

724-357-2956

Fax: 724-357-4842

www.iup.edu/sociology/als

Informed Consent Signature Page

VOLUNTARY CONSENT FORM:

I have read and understand the information on the form and I consent to volunteer to be a participant in this study. I understand that my responses are completely confidential and that I have the right to withdraw at any time. I have received an unsigned copy of this informed Consent Form to keep in my possession.

Name: (PLEASE PRINT) _____

Signature: _____

Date: _____

Phone number or email where you can be reached: _____

Best days and times to reach you:

Appendix C
IRB Approval Letter



Indiana University of Pennsylvania

www.iup.edu

Institutional Review Board for the
Protection of Human Subjects
School of Graduate Studies and Research
Stright Hall, Room 113
210 South Tenth Street
Indiana, Pennsylvania 15705-1048

P 724-357-7730
F 724-357-2715
irb-research@iup.edu
www.iup.edu/irb

October 31, 2016

Stacey Tice
503 South Juliana Street
Bedford, PA 15522

Dear Ms. Tice:

Your proposed research project, "Staff Perceptions of Impact of Performance Measurement Requirements on Program Delivery and Organizational Effectiveness in Nonprofit Human Service Agencies," (Log No. 16-243) has been reviewed by the IRB and is approved. In accordance with 45CFR46.101 and IUP Policy, your project is exempt from continuing review. This approval does not supersede or obviate compliance with any other University requirements, including, but not limited to, enrollment, degree completion deadlines, topic approval, and conduct of university-affiliated activities.

You should read all of this letter, as it contains important information about conducting your study.

Now that your project has been approved by the IRB, there are elements of the Federal Regulations to which you must attend. IUP adheres to these regulations strictly:

1. You must conduct your study exactly as it was approved by the IRB.
2. Any additions or changes in procedures must be approved by the IRB before they are implemented.
3. You must notify the IRB promptly of any events that affect the safety or well-being of subjects.
4. You must notify the IRB promptly of any modifications of your study or other responses that are necessitated by any events reported in items 2 or 3.

The IRB may review or audit your project at random or for cause. In accordance with IUP Policy and Federal Regulation (45CFR46.113), the Board may suspend or terminate your project if your project has not been conducted as approved or if other difficulties are detected

Although your human subjects review process is complete, the School of Graduate Studies and Research requires submission and approval of a Research

IRB to Stacey Tice, October 31, 2016

Topic Approval Form (RTAF) before you can begin your research. If you have not yet submitted your RTAF, the form can be found at <http://www.iup.edu/page.aspx?id=91683>.

While not under the purview of the IRB, researchers are responsible for adhering to US copyright law when using existing scales, survey items, or other works in the conduct of research. Information regarding copyright law and compliance at IUP, including links to sample permission request letters, can be found at <http://www.iup.edu/page.aspx?id=165526>.

I wish you success as you pursue this important endeavor.

Sincerely,

Jennifer Roberts, Ph.D.
Chairperson, Institutional Review Board for the Protection of Human Subjects
Professor of Criminology

JLR:jeb

Cc: Dr. J. Beth Mabry, Dissertation Advisor
Dr. John Anderson, Graduate Coordinator
Ms. Brenda Boal, Secretary

Appendix D

Coding and Analysis

This appendix provides a detailed description of the process I used for coding and analysis to allow other researcher to replicate my process and validate my findings. This study involved complex relationships between the organizational effectiveness and staff member perception and several steps where involved to allow me to view the data with consideration of the study participants place within the agency and patterns or trends among the agencies. This was a qualitative study with a small sample size, 28 interviews in seven agencies. Those characteristics were helpful in understandings context of the responses and provided insight to the overall patterns using an interpretive style of analysis (Welsh, 2002). I did not intend to have findings that would be generalized about the staff member position within the agency or agency characteristics.

This study was about staff perceptions of performance measurement. The purpose of the study was to understand how these perceptions could inform researchers and policy makers of the impact of performance measurement on organizational effectiveness. It also provided some insight into how performance measurement occurs within the agencies and the ability of performance measurement to be used for its intended purposes: resource allocation, program improvement, and accountability. Based on existing literature on organizational effectiveness and performance measurement, I developed three research questions as follows:

1. What is the impact of performance measurement on service delivery and organizational performance relative to the mission?
2. Do NPO staff members feel they have adequate (a) training and (b) time to collect quality performance measurement data?
3. How are performance measurement results communicated to caseworkers for improving the delivery of programs and services?

I developed the following interview questions to allow study participants to speak to the research questions:

1. As part of your job, I understand that you collect information about your employment and training clients and programs. What's your experience with that?
 - a. Tell me about the types of information and how you collect it?
 - b. Do you run into any difficulties gathering the necessary data?
2. How does collecting this information affect how you do your job?
 - a. Tell me about the time it takes?
 - b. How does that impact the help you can give clients?
3. What's your sense of how this information gets used in terms of programs for service delivery to clients?
 - a. What types of reports are generated from the information?
 - b. How is funding affected?
4. Tell me about the training you get for collecting and using this information?
 - a. What types of trainings do you wish you received?
5. Tell me about how you are involved in the process; what input do you have?
 - a. Do you track more information than is required?
 - b. From your perspective, what would it take to make the information you collect more meaningful?
6. What type of feedback do you get from the information you collect?
7. What type of feedback would you like to get?

Because the interviews were done in a conversational style, the answers were often provided within a response to another question. I also found that many of the study participants wanted to talk about the details of the program and their work experiences. This provided rich context for their responses. The challenge of this interview style was that while study participants were comfortable in speaking about their experiences; the information was not easily ordered in the interviews. To address this challenge, I conducted the coding in three steps.

The first step required a thorough reading of the transcripts while re-listening to them. I coded the responses by chunking text as it related to the research questions throughout the

interviews, not strictly by questions. This process provided some insight into the staff member perceptions, but did not provide enough detail to allow for analysis. Several broad themes did emerge that spoke to the research questions, 1) the disconnect of the performance measures to program activities, 2) the difficulties that occur when performance measurement is used for many purposes and paired with other processes, and 3) the external influences on performance measurement. To get a more thorough and complete understanding of the interview material, I coded the material again.

For this second step of the coding process, I reviewed the transcripts by interview questions. Since the interviews were done in a conversational style, I reread all of the transcripts for each question to make sure that no response related to the questions was missed. Using Microsoft Excel software, I created a template of study participants that was color coded to show the agency and their role within the agency as either a caseworker, a manager or a staff member that had dual responsibilities. This provided a multidimensional visualization of the material. I reviewed each question for the major elements of that question, i.e. training, feedback, impact, type of information, feedback, reports, etc. And I also searched for comments that related to the elements of the first review which were influences, disconnect and the differing processes within and between agencies. The text of the responses that related to those elements were pasted into the spreadsheet in a row for that study participant with a sheet for each element. This process allowed me to add rows for multiple comments about a single element from a study participant. This resulted in very large spreadsheets that allowed me to visualize the information according to the major elements of my questions and the agency and position of the person making each response. This process also allowed me to see other concepts that were not actually an element of the questions. An example is the frequency that staff members talked about the importance of

having a relationship with the clients. I reviewed the entirety of the transcripts again to be certain those types of comments were not missed after all the initial questions were coded. This provided me with an analytic structure that I used to identify the direction and frequency of related responses in each category “across and within groups of questions” (Bradley, Curry & Devers, 2007, p. 120) as well as across and within groups of study participants.

After I was confident that all of the interviews had been thoroughly mined for information related to my research questions, I began to analyze the information on each spreadsheet to determine patterns and categories. Additional columns were added to the spreadsheet according to these categories and each response was determined to be in a category or not. This was an iterative process since as the responses were reviewed, the categories became more refined to accommodate multiple related comments.

In some instances, such as with feedback, the responses from some participants seemed contradictory. When this occurred, I went back to the transcripts to review the context of the response. This allowed me to reexamine my categories to provide a more complete and in-depth analysis of the responses and not simply good, bad, or neutral.

Once I was satisfied with my coding and review of the responses, I wanted to provide some perspective relative to frequency of responses. I also wanted to provide some context for responses related to frequency relative to position within the agency when that information was relevant. To do this I created columns to count responses by position within the agency or by the entire number of study participants. This also expanded the analysis respective to their position within the agency as being a caseworker, manager or having dual responsibilities.

The findings of this study were based on the patterns and themes that I detected through this process. This process allowed me to provide detail, context and an objective representation of the study participants views.

REFERENCES

- Bradley, E. H., Curry, L. A., & Devers, K. J. (2007). Qualitative data analysis for health services research: developing taxonomy, themes, and theory. *Health Services Research, 42*(4), 1758-1772.
- Welsh, E. (2002). Dealing with Data: Using NVivo in the Qualitative Data Analysis Process [12 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 3*(2), Art. 26, Retrieved from <http://www.qualitative-research.net/index.php/fqs/article/view/865/1880>.