5-1-2021

Nonprofit Data Management: A Stage Model

Ashley M. Hernandez-Hall

NONPROFIT DATA MANAGEMENT: A STAGE MODEL

By

Ashley M. Hernandez-Hall

Bachelor of Social Work
University of Nevada, Las Vegas
2011

Master of Social Work
University of Nevada, Las Vegas
2014

A dissertation submitted in partial fulfillment
of the requirements for the

Doctor of Philosophy – Public Affairs

School of Public Policy and Leadership
Greenspun College of Urban Affairs
The Graduate College

University of Nevada, Las Vegas
May 2021

Dissertation Approval

The Graduate College
The University of Nevada, Las Vegas

February 5, 2021

This dissertation prepared by

Ashley M. Hernandez-Hall

entitled

Nonprofit Data Management: A Stage Model

is approved in partial fulfillment of the requirements for the degree of

Doctor of Philosophy – Public Affairs
School of Public Policy and Leadership

Patricia Cook-Craig, Ph.D.
Examination Committee Chair

Christopher Stream, Ph.D.
Examination Committee Member

Jessica Word, Ph.D.
Examination Committee Member

Jennifer (Starr) Hoffman, Ph.D.
Graduate College Faculty Representative

Kathryn Hausbeck Korgan, Ph.D.
Graduate College Dean

Abstract

Managing nonprofit data is both complicated and essential. Nonprofits, while struggling to manage and utilize data to its fullest potential, are increasingly required to do so for funding purposes. Despite the increased pressure on nonprofit managers to report more outcomes to grant funders, little research is available to guide the data management process in nonprofits. Research that does exist is primarily focused on the for-profit or business sector, which is operationally and fundamentally different than nonprofits. For example, for-profit entities typically do not have the same restrictions on how to use funds, such as a percentage cap on spending for non-direct business costs (or overhead), that nonprofits must contend with. Additional funding restrictions, such as funders not allowing infrastructure spending, further constrain how nonprofits manage their technology and their data. As such, the research, and recommendations for data management in the for-profit sector are often not as applicable to the nonprofit sector.

This dissertation sought to discover the ways in which nonprofits manage their data, and whether those data management practices are related to nonprofit program outputs and outcomes. Utilizing Stage Theory, the literature review focused on the data collected and used by nonprofits, as well as their data management practices, and that review was used to create a Nonprofit Data Management Stage Model. The model organizes nonprofit data management practices into four separate practice domains and places data management elements in each domain into five distinct stages of data management. Using results from a survey sent to domestic violence shelters across the nation, reported data management practices were placed within the various domains and stages within the Nonprofit Data Management Stage Model. Output and outcome measures collected from the same survey were used to test the relationship between data management practices and program outcomes of nonprofits.

Acknowledgements

First, to Dr. Patricia Cook-Craig. Thank you for sticking with me, helping me get my dream job, and figuring out a way to keep my data voice in check while writing. I could not ask for a better mentor, especially during the wild ride that was 2020.

To the UNLV Public Policy and Leadership faculty, and especially my dissertation committee, I thank you. You helped me grow and you gave me the confidence to shine. I feel successful, which is such an amazing gift. I carry your example and your expectations with me.

Dr. Leroy Pelton, I am sad you are no longer here as I would have loved to share this accomplishment with you. You encouraged me to tackle graduate school and pushed me on a path I was not expecting. Your influence and passion are forever a part of who I have grown to be. I hope you knew how special you were.

I want to thank all the people who helped me through some of the most tedious parts of this journey. Stormy and Landon, Aunt Diana, and of course Mom, I couldn't have done it without you.

To my parents, who never held me back or asked too much of me. You let me grow into what I was to be, and you encouraged me along the way. I owe my independence to you.

I want to thank all of my siblings; you have believed in me since I was a nerd in grade school. Thank you. I won't let you down.

Thank you to Grandma and Grandpa Hill; you expected this of me. And thank you Lynn Hall and Gene Hall. You became my parents through marriage when I was already grown, and yet you still showed a pride in me as if I was one of your own children.

I want to thank my kids. You may not have known all the details or how long this haul really was, but I can tell you are proud of me. That means the world.

And finally, thank you Vince. You endured my endless brainstorming, planning, frustrations, and writing sessions. You were a sounding board, you proofed my work, and you slogged through right alongside me. You deserve this as much as I do. I can't imagine that I would have completed this journey without your constant support.

Dedication

I dedicate this, my final act as a student, to every first-generation college student who doesn’t think they are meant for higher education. You are meant for what you desire, nothing more, nothing less. You belong and you are worthy. Go for it.

Table of Contents

List of Tables

Chapter 1: Introduction

Introduction

Data management is essential to any organization that wishes to collect, store, and utilize data for planning, evaluation, and reporting. This is especially true for nonprofit organizations who receive grant funding from various private and public sources, as those funds usually have reporting requirements attached to them. However, the data itself is generally misunderstood by nonprofit managers. This is a challenge for nonprofits as data that is unmanaged or mismanaged may lead to poor decision making, inaccurate statistical reporting, and inaccurate program or organizational evaluation. Without accurate data, it is difficult to get a clear picture of the health of an organization, or the realistic impact of the services that nonprofit organizations provide to clients and communities.

While the data itself is misunderstood in the nonprofit world, the data management practices of nonprofits are even more so. Although ample research exists on the various nonprofit activities that are occurring which utilize data, such as program evaluation, needs assessments, or survey methodology, very few researchers have studied the actual data management practices of nonprofit organizations. While limited, the research that does exist points to data being used for a variety of reasons. Service improvement activities, including program evaluation and satisfaction studies, are common data-heavy nonprofit activities according to multiple studies (Botcheva, White, & Huffman, 2002; Carman & Fredericks, 2008). Program planning, including needs assessments and strategic planning, which help "establish program goals or targets," is also frequently practiced by nonprofit organizations (Carman & Fredericks, 2008, p. 58). One of the most frequent uses of data by nonprofits, however, is reporting to program funders, which is essential and often a primary focus, as funding is necessary for service delivery (Botcheva, White, & Huffman, 2002, p. 428). Other less frequent uses of data include resource planning and funding diversification (Botcheva et al., 2002; Carman & Fredericks, 2008; Lenczner & Phillips, 2012).

This lack of understanding of requisite data management skills, leadership, and organizational cultural supports required for a successful data-driven nonprofit, all contribute to significant challenges with the data itself. Ineffective management, including data quality assurance and security problems, lead to added costs for organizations such as staffing time spent on correcting errors and reduced program effectiveness, especially if faulty data is used to start new programs (Redman, 1998, p. 80). In addition, many missed opportunities can arise when data is not managed effectively. For example, quality data, analyzed and presented in a concise way, could be used to reduce programmatic costs, improve decision making, improve operations, and increase interdepartmental or even interorganizational data sharing and program collaboration (Beath, Becerra-Fernandez, Ross, & Short, 2012; Gamage, 2016).

In short, poorly managed data can lead to costly mistakes, missed opportunities, and faulty reporting, all challenges that can reduce the positive impact nonprofit organizations hope to achieve for their client populations.

Presently there is no unified model or theory of nonprofit data management which can guide both the study of data management in the nonprofit sector, or the activities of nonprofits to help them become more efficient in their data management practices. This dissertation sought to understand and explain the concept of data management and its history, as well as the data collection and use practices of nonprofit organizations. By examining current literature on the topic of nonprofit data use and practices, and utilizing Stage Theory as a theoretical framework, a nonprofit data management model was developed and presented, which identifies the various data management activities and how they can evolve over time as an organization's data management practices mature. This data management framework was then used to develop a survey to collect information from nonprofit organizations to better understand their current data management practices. The information collected was then used to explore the relationships between nonprofit data management practices and their program outputs and outcomes.

History of Data Management

Data is often misunderstood to mean numbers, usually financial in nature, that are stored in a spreadsheet or database on a computer. However, data encompasses a range of informational inputs that have existed long before the computer was invented. Gray (1996) synthesized the history of data management into six separate time periods that track data management practices from 4000BC to today (p. 2-8). The Zeroth generation included record managers between 4000BC to the 1900s and includes data, typically tax and/or census data, recorded in clay tablets, parchment, and eventually paper. The First generation, existing between 1900 and 1955, saw the rise of automated information processing including punch cards created in fabric which eventually progressed to paper. While the first computers were being developed during this period, it wasn't until the Second generation, between 1955 and 1970, that those computers were used for data storage and analysis. These computers were initially used for performing calculations, and the data developed during this process was secondary; however, as technology evolved to include magnetic tape that could store information for future retrieval, the concept of data storage began to take hold. It was during this second generation that software was developed to capture and process stored information, and as computing prices dropped, hardware used to analyze data was purchased by smaller and smaller businesses. As technology entered the Third generation, from 1965 to 1980, computer systems were being used to handle stock-market trading and data was moved to an online system as software became capable of handling "concurrent transactions against a database shared among many terminal users" (Gray, 1996, p. 5). The Fourth generation, which occurred between 1980 and 1995, moved data analysts and managers forward with the advent of computer programming using relational models, which manage data using a "unified language for data definition, data navigation, and data manipulation, rather than separate languages for each task" (Gray, 1996, p. 6). This paved the way for structured query language (SQL) and graphical user interfaces (GUIs), which allow users to "pose complex database queries" and output data into readable, manipulatable tables and graphics. While much of the technology and tasks that developed in the Fourth generation are still being used today, the Fifth generation, which represents the time between 1995 and present day, has seen a rise in technology that is capable of processing, storing, and searching data that isn't simply numbers or sets of records. Complex data objects, such as images, maps, or sound files, as well as massive "big data" sets are being stored in multimedia databases, where it is being used by everyone from web surfers to researchers, to law enforcement and policy makers.

While this brief history of the use of data may make it appear as if data management has progressed smoothly to the usable format that we see today, the reality is that data storage and use is incredibly misunderstood by all but the most experienced programmers and data scientists. In fact, it is often the case that data users in larger organizations do not know where or how data is stored or processed into the usable format that they enjoy. Currently, most organizations use only data that they themselves create, and smaller organizations, such as small nonprofits, rarely, if ever, use complex databases. However, even low-volume simplistic data needs to be managed to ensure accuracy, security, and usability.

Problem Statement

Many studies have outlined the type of data collected in nonprofits, for example, NTEN (2012) reports that 99% of the nonprofits surveyed for their study collected some sort of metrics, including financial, outcome, donor, and external data. The uses of data, especially in evaluation activities in nonprofit organizations, is well represented in academic literature. However, while evaluation data may be used frequently, there is numerous data of various types that go unused due to lack of knowledge or experience (Gregory & Howard, 2009; Mitchell & Berlan, 2016; Stoecker, 2007), lack of staff or resources (Gregory & Howard, 2009; Wing, 2004), or due to a lack of an organizational culture that is consistent with data use (Botcheva, White, & Huffman, 2002; Carman & Fredericks, 2008; Mitchell & Berlan, 2016). And while a significant amount of research exists about nonprofit evaluation and outcomes activities, including how to conduct evaluation studies, there is little research that is focused on managing the data created or collected to perform those studies. This is problematic as funders of nonprofit programs often require data reports including client outcome reports, output statistics, and program evaluations. However, without the necessary funding, resources, experience, expertise, or organizational capacity to collect, store, and process various data, accomplishing the required evaluation and outcomes activities is challenging at best.

Beyond evaluation, nonprofits perform additional tasks that require the use of data. For example, nonprofits often conduct vital research, including needs assessments, policy analyses, and academic research in their field of expertise, and data is necessary to those research studies. However, restraints and burdensome requirements put on nonprofit organizations by funders, specifically the need to provide output reports, as well as the lack of flexibility to utilize funding for data technology (such as databases or information systems) and data experts, all hamper the ability of nonprofits to effectively use data (Ebrahim, 2002, p. 108-110). Even if restrictive funding was not an issue, nonprofits still have the problem of data management experience to tackle. Stoecker (2007) found that while nonprofits collect a large amount of data, and spend significant manhours doing so, the data is rarely used (p. 108). This is likely due to the fact that nonprofits often have very little research methods experience, making the management and use of data a difficult skill to master for the average nonprofit employee or manager (Stoecker, 2007, p. 111-112).

This ineffective use of data leads to a variety of problems. Aside from the obvious inaccurate or missing reports, data management challenges could represent missed opportunities for nonprofit organizations, which can lead to an ineffective use of public and private funding. Beath et al. (2012) explain that the use of the expanding data collected by organizations can bring benefits such as cost reductions, improved research speed, improved operations, and in the case of data sharing, enhanced interorganizational collaboration (p.19-20). Gamage (2016) echoes these findings by explaining that big data specifically can be harnessed by governments to improve service delivery, such as social service and transportation, as well as improve decision making around public health, safety, defense, and disaster management (p. 386). And while these missed opportunities are significant, poor data management practices may actually end up costing nonprofit organizations money in the long run. Redman (1998) explains that consequences of data management and data quality failures include increased operational costs due to finding and correcting errors, reduced trust in organizations both by customers and clients, and by employees as data errors become apparent, and reduced effectiveness of programs and strategies developed with faulty data (p. 80).

While well-managed data can be an invaluable resource to organizations that wish to strategically plan, accurately evaluate, begin interorganizational collaboration, or improve operations, conversely, data that is not well managed can have the negative effect of draining man-hours from organizations, harming reputations and eroding trust. While funding restrictions are a major problem for nonprofit organizations who wish to strategically manage data, additional concerns, including a lack of experienced personnel or lack of organizational commitment, can lead to a poor data management strategy or no data management strategy at all.

Research Purpose

The purpose of this research is to explore the relationship between nonprofit data management practices and program outputs and outcomes. Little research has been done in this area, and as such this study was primarily an exploratory analysis of the data management practices of nonprofit organizations. Using Stage Theory and a thorough literature review, a five-stage nonprofit data management model was developed which describes the various data management practices that could be present within a nonprofit organization. In each of the five stages, nonprofit data practices were broken down further into four data management domains, including culture, leadership and planning, procedures, and budget. A 50-question survey was designed and used to explore the data management practices of one type of nonprofit organization, domestic violence (DV) shelters. The results of the survey were used to explore the relationships between data management practice of domestic violence shelters to selected DV shelter outputs and outcomes.

Research Questions

This dissertation sought to answer the following question: are data management practices related to nonprofit program outputs and outcomes? To answer this question, the following three hypotheses were tested:

> Hypothesis 1: Nonprofits at a later stage of nonprofit data management will have increased outputs compared to those at an earlier stage of nonprofit data management.
>
> Hypothesis 2: Nonprofits at a later stage of nonprofit data management will use data for decision making more than those at an earlier stage of nonprofit data management.
>
> Hypothesis 3: Nonprofits at a later stage of nonprofit data management will have better program outcomes than those at an earlier stage of nonprofit data management.

Significance of the Study

This research sought to fill a substantial information gap. Little existing research can be used to develop a unifying theory of nonprofit data management, and because of that, very little prescriptive research exists that can guide nonprofit managers in the realm of data management. While ample research is available to help guide nonprofit managers in various data-heavy activities, such as program evaluation or conducting needs assessments, there is still much effort and struggle to manage the data that is created and stored. In addition, by limiting data activities to these isolated tasks, the full value of that collected data is never realized. By broadening our understanding of the types of data collected by nonprofits, and the activities that nonprofits are currently undertaking in the realm of data management, a larger body of research can be tapped to develop models and frameworks to guide nonprofit data management and use.

Definition of Terms

A variety of terms were used in this dissertation to describe data management practices and the proposed stage model. The most significant and prominent of these are listed below, with accompanying definitions.

- Data management. The various tasks involved in creating, collecting, storing, using, protecting, and sharing information.
- Data management domain. A collection of practices that make up a common theme or practice area of data management. For this research, four data domains will be discussed: culture, leadership and planning, procedures, and budget.
- Data management element. A data management element is a specific task or action that is performed during data management. Multiple data elements make up a data management domain. For example, one data element in the budget domain is an organization paying for data training for staff.

Delimitations, Assumptions, and Limitations

Delimitations

This research sought to understand the data management practices of nonprofit organizations and how they relate to program outcomes; as such, only nonprofit organizations were selected as a study population. Because nonprofits are incredibly diverse in their operations, services, practice areas, locations, and the populations they serve, additional narrowing was necessary. In order to utilize a population that was both widespread (throughout the United States) and consistent in their practice and the populations served, domestic violence shelters providing emergency shelter services to clients were selected.

In order to collect data that was consistent from respondent to respondent, a survey instrument was developed that only allowed selection of pre-designed options. No qualitative data was utilized in this analysis, though data of that nature might be useful in broadening the understanding of data management practice of nonprofits in future research.

Assumptions

Since a survey instrument was developed to collect information about the data management practice of the nonprofit organizations sampled, it was assumed that survey participants were knowledgeable about the data management practices of the organization for which they responded. Every effort was made to ensure questions were worded for maximum understanding of data management practices, including reducing or eliminating nomenclature that might not be fully understood by those inexperienced in data management.

In addition, it was assumed that all respondents worked in nonprofits that did collect, store, and use data. This was a fairly safe assumption, however, as domestic violence shelters rely heavily on governmental funding, which requires data collection and reporting in order to qualify for funding.

Because research in this area is limited, this dissertation is exploratory in nature. While a thorough literature review was conducted to help guide the development of a data management stage theory and a conceptual model, much of the elements and organization of those models were developed using previous research in areas outside the nonprofit sector, or were extrapolated from research that had a very narrow focus (such as data technology spending, for example). Therefore, the final Nonprofit Data Management Stage Model is based primarily on a literature review of incomplete literature.

Limitations

As with any exploratory study, limitations are present in this dissertation. First, and foremost, the survey itself has limitations. There was no survey instrument in existence that could collect the information needed to learn about nonprofit data management practices, so one was developed. The scope of this research did not allow for an instrument to be piloted and tested, so the instrument's reliability is a limitation. In addition, the survey was delivered via email as an online survey which typically results in a low response rate, as indeed was the case. Another limitation was survey length; while an attempt was made to keep the survey as short as possible to increase completion rates, this meant sacrificing some questions that might have otherwise been asked. For example, a limited number of demographic questions were asked about the survey taker, and few questions were asked about the type of clients the organization serves.

Conclusion

This dissertation was an exploratory analysis of the data management practices of domestic violence emergency shelters. By exploring the current practices and perceptions around data management, as well as reviewing the literature on the nature of nonprofit data management, this study developed and presented a stage model of nonprofit data management. Using the data collected from an online survey, organizations were placed into a stage on the stage model and then the relationships between nonprofit data management maturity and program outcomes and outputs were explored.

Chapter 2: Review of the Literature

Introduction

In this chapter, a thorough literature review that is organized by data management themes is presented and the literature review process is described. Next the data management domains, culture, leadership and planning, procedures, and budget, are introduced. Finally, Stage Theory is described and the Nonprofit Data Management Stage Model and research conceptual model are presented.

Reference Lists

Due to the limited information on nonprofit data management practices, the literature review for this dissertation was intentionally thorough and followed several steps. A list of search terms was compiled including nonprofit data management, data management, nonprofit evaluation, nonprofit big data, data management, big data, assessment culture, nonprofit assessment, data management budget, and stage theory. The resulting articles and their reference list were reviewed. Once this process was complete, a list of 129 sources was collected for review.

In order to ensure other relevant sources were not overlooked, a review of all reference lists for these sources was conducted. A dataset was created in Microsoft Excel that included a list of every article or book cited in each of the original sources. The completed dataset contained 3,326 entries, though this included duplicates (some titles were listed in multiple articles). Once this dataset was complete, Tableau (a data visualization application) was used to count the number of times each title appeared in the dataset. Most titles appeared only once in the list and were eliminated. Eighteen titles appeared four or more times, meaning they were cited by four or more of the initial literature review sources. Of these 18 titles, four were found to be inappropriate for inclusion in this study, four were not able to be located, five were already included in the original source list, and five were added to the literature review, bringing the total number of sources used in this research to 134.

Literature Review

Reviewing the literature revealed a significant lack of research on how nonprofit organizations utilize data. Most of the literature was aimed at the private sector, specifically larger organizations with heavy data use needs, and emphasized the IT and/or server management practices rather than the human and organizational culture aspects of managing data. Research that focused on nonprofits typically focused on the evaluation, program management, or research practices of nonprofits. In spite of the fact that data is required to carry out evaluation functions, the management of that data is often ignored or only briefly mentioned in the context of organizational challenges.

While data management literature is fragmented, the review did illuminate certain themes. The literature review findings were placed into the following data management domains that illustrate these themes: culture, leadership and planning, procedures, and budget. Within each of these domains, researchers explored various occurring practices, challenges, and recommendations related to data management in the public, private, and nonprofit sectors. In addition to the data management practices that were identified, the literature also revealed the data types and ways in which data are used in the nonprofit sector. The following literature review is organized into two main sections: Nonprofit Data Types and Uses, and Data Management Practices and Recommendations.

Nonprofit Data Types and Uses

Data Types

Data can be broken into hundreds of categories depending on its format, uses, creation method, location, and a variety of other characteristics. However, defining data by such small parameters is often not useful. As such, this research focused mainly on the overarching data storage method, as this will have direct implications on how that data is managed and utilized in an organization. The three types of data discussed in this research are digital data, paper files, and big data.

Digital Data. This is data, according to Lenczner & Phillips (2012), that is in a digital (computerized) format that can be read and interpreted by a computer program. Examples of this include social media posts or hits, website traffic, or digital documents, such as PDFs or Word documents. Files that are created by nonprofits in a digital format, such as online intake forms or digital case notes, are common examples of digital data in the nonprofit environment.

Paper Files. Data found in paper files are still prevalent in nonprofit organizations. Stoecker (2007) found in a survey of 80 nonprofit organizations that "on average, 61% of the data [collected by nonprofits] is saved in paper files" (p.108). Carman and Fredericks (2008) also found that paper data was widespread in their survey of 340 nonprofits. They found that 79% of their sample relied on "written data collection tools" for their evaluation activities (p. 58). This paper data can include anything from case notes, to intake forms, to satisfaction surveys.

Big Data. Big data is a term used often in literature without providing a clear definition. However, De Mauro, Greco, and Grimaldi (2016) analyzed the literature to synthesize the various themes associated with big data into a unified definition: "Big Data is the information asset characterized by such a high volume, velocity and variety to require specific technology and analytical methods for its transformation into value" (p. 131). In other words, big data specifically refers to data that is so large, so varied, and that moves and changes so quickly, that it requires specialized technology to analyze it. While not specified in the above definition, big data is digital data. While this does not seem like a data type that would be used by resource limited nonprofits, the exploration of big data is gaining traction, and its uses are varied.

Data Uses

The ways in which organizations use data is even more varied than the types of data that exist. Again, breaking down each individual use for data would be unhelpful to the current research, so the uses of data is organized into themes: program planning, service improvement, external reporting, and other data uses.

Program Planning. Carman and Fredericks (2008) explain that 75% of the 340 survey respondents in their study reported using data to help “establish program goals or targets” (p. 58) and another 68% reporting using data to develop new programs. Botcheva, White and Huffman (2002) came to a similar conclusion, reporting that 71% of their research participants utilized collected data for strategic planning while 75% used data to assess program implementation progress (p. 428). Gamage (2016) found that public organizations could utilize big data, specifically, to “make better informed policy decisions and to address the citizens’ needs more appropriately” (p. 386).

Service Improvement, Evaluation, and Outcomes Measurement. Service improvement and evaluation is likely one of the first data uses that comes to mind when considering nonprofit data use. Examples of these types of data task include program evaluation, user satisfaction surveys, or program completion statistics. According to Carman and Fredericks (2008), 93% of surveyed nonprofits used data to change existing programs, the most common use of data for organizations in the study. Botcheva, White and Huffman (2002) also found that a significant percent (75%) of organizations used collected data for service improvement while another 42% of organizations utilized data to measure client satisfaction (p. 428).

External Reporting. Botcheva, White and Huffman (2002) found that the most common use of data was to report to funders: 88% of participants reported using data for this task (p. 428). Carman and Fredericks (2008) reported a slightly lower percentage of organizations using data for funder reports, with only 67% reporting use of data for this task. However, they also found that 82% of organizations used data for reports to boards of directors.

Other Data Uses. Carman and Fredericks (2008) found that very few organizations utilized data for tasks such as outreach or to find additional funding. Botcheva, White and Huffman's (2002) survey results supported these findings, with only 8% of their participants reporting using data for unspecified "other" reasons (p. 428). These findings suggest an underutilization of data, an idea that is echoed by Lenczner and Phillips (2012) who explained that "the clearest opportunity [for data mining] is to use this information to benefit resource planning, specifically the search for diversified, stable funding" (p. 13).

Data Management Practices and Recommendations

Culture

While data management is often viewed as a purely technical function, the reality is that for a data management plan to be successful in an organization, the culture of that organization needs to be amenable to it. Weber and Kristin (2007) explain that for a successful model of data governance, organizations need to recognize that information technology (IT) governance and data governance are coequal activities that must follow an organization-wide governance principle. This concept of collaboration is essential, and goes beyond the IT department and upper management, as data management should be an activity that is shared among multiple levels of an organization (Thompson, Ravindran, & Nicosia, 2015). In order to create and maintain a management plan that succeeds and is representing all organizational needs, the plan for managing and using data needs to be centralized and holistic; organizations must commit to treating data as a resource and must identify how that resource can provide value to the organization (Beath, Becerra-Fernandez, Ross, & Short, 2012; Cong & Pandya, 2011; Khatri & Brown, 2010; Levitin & Redman, 1998; Malik, 2013). Gantz and Reinsel (2011) agree, explaining that in order to gain value from the varied data that an organization collects and stores, the organizational culture must value data-driven practice and decision making. Botcheva, White and Huffman (2002) also found that the consistency of data collection efforts was positively correlated with the learning culture of an organization (p. 430). Similarly, Mitchell and Berlan (2016) found that “a desire to improve program effectiveness” was the most important catalyst to evaluation. However, without a strong data management culture, that desire to improve cannot be cultivated.

Leadership and planning

Leaders often drive the important decisions that contribute to a functional data management strategy. In their qualitative case study, Thompson, Ravindran, and Nicosia (2015) found that leadership is “crucial for ensuring the success of any data governance initiative” (p. 320). They go on to state that “an effective leader, who clearly communicates and directs the direction of the organization can accomplish things that would be impossible otherwise” (p. 320). In literature focused on academic libraries, research on the culture of assessment (evaluative practice that often involves data) consistently shows that strong leadership, or leadership that is dedicated to making assessment a part of daily practice, is a necessary component to its success (Farkas, Hinchliffe, & Houk, 2015; Lakos & Phipps, 2004; Ndoye, 2010).

In business literature, Beath et al. (2012) found that the explosion of data collected by organizations, both structured and unstructured, has created significant challenges in the area of data responsibility; even identifying who should take the lead with data management tasks does not occur regularly (p. 18-19). This finding is similar to those described by Gantz and Reinsel (2011). They explain that during their research, the most surprising challenge to data management was “the cultural challenge” (p. 8). They further explain that while big data projects should be prioritized and approached strategically, organizations were instead treating them as “junior science projects” with limited assets or experts dedicated to them (p. 8). Levitin and Redman (1998) explain that due to the significant challenges associated with managing data effectively in an organization, it is necessary for senior executives, as opposed to lower-level managers, or departmental heads, to lead the data management process. They warn that “without strong leadership from the top, data management programs risk falling prey to the disparate agendas of functional areas and their leaders” (p. 100).

Even for the leader who takes data management seriously, finding staff to accomplish the tasks is still a significant challenge. Carman and Fredericks (2010) found that challenges associated with evaluation activities were related to organizational staff's dissatisfaction with their own technical abilities. According to Lee and Lan (2011) “appropriate support and training programs” are “imperative” to the adoption of a new knowledge management system (p. 733). Bernard and Pukstas (2009) showed that those organizations who considered themselves to be technological leaders generally had a formal (strategic) plan for the use of technology, and those organizations were also more satisfied with IT functions than those that did not consider themselves technological leaders. This may be due, in part, to the fact that technology leaders were more likely to provide training which led to staff "who are well-trained and comfortable with technology [and] are better prepared to meet an organization's technology needs" (Bernard & Pukstas, 2010, p. 23). However, few nonprofit organizations committed to IT training. Echoing this finding, Stoecker (2007) reports that the number of staff or volunteers in nonprofit organizations surveyed that were formally trained in research averaged less than one person per organization, and nearly half reported having no one at all on staff with formal research training (p. 111). Gamage (2016) also reported a skills shortage in data analytics in the public sector, which leads to challenges in data management, especially in the emerging big data analytics field (p. 388). Cong and Pandya (2003) described an environment in which organizations which attempt to manage not only the data, but the knowledge of an organization, run into roadblocks when staff simply do not understand what knowledge management is or how it can help. To mitigate this challenge, Gantz and Reinsel (2011) suggest that it is important for leadership to lay the groundwork for valuing data. This can be done through cultivating skills and interest within the organization. Leadership must make skills cultivation a priority, either through hiring skilled staff or training existing staff, for data management strategies to succeed.

While necessary research and data skills are needed, role clarification is also important. Data ownership role expectations, including managing the collaboration between IT and data users, as well as identifying explicit accountability requirements, is necessary for leadership to manage data effectively (Beath et al., 2012; Khatri & Brown, 2010; Levitin & Redman, 1998; Malik, 2013; Weber, Otto, & Österle, 2009).

Beyond staff skills and role clarification, leadership must recognize the importance of properly managed data for decision making. Redman (1998) explains that decisions made based on poor data can lead to significant problems as “decisions are no better than the data on which they are based” (p. 81). Poor quality decisions can also lead to mistrust in the organization or leadership, which can further erode the data management culture of an organization. However, leadership that prioritizes a data plan and works to ensure appropriate communication between data managers and business leaders, will have a better chance of success in their work (Thompson et al., 2015, p. 320).

Nonprofit leaders are also essential because they typically initiate planning. Planning for data use is important but should go beyond simply listing data that needs to be collected and how it should be stored. For a properly managed data system, data use and management should tie directly into organizational goals and principles (Khatri & Brown, 2010; Malik, 2013). However, research has shown that a significant portion of data collected by nonprofit organizations is not data that is internally motivated and is not used beyond externally required reporting. Stoecker (2007) found that while half of the participants in their study were required to collect data by funders for evaluation, that data was not used beyond those evaluations. In other words, the organizations found that “things they are required to report on do not help them actually do their own work” (p. 109).

Externally motivated data collection and reporting, or data reporting often required by external funders only, can lead to significant challenges in the development of a data-driven organization. For example, Carnochan, Samples, Myers, and Austin (2013) found that data management suffered in organizations in which management and staff felt that evaluation activities were arbitrary or unimportant. Conversely, research has shown that organizations that link evaluation and performance measurement with broader organizational goals, or those that are motivated internally to evaluate, exhibit more rigorous evaluation activities (Carman & Fredericks, 2010; Mitchell & Berlan, 2016).

Once data-heavy activities, such as evaluation, are completed, it is important for those reports to be shared so that data collectors, analysts, and reporters can see the value and use of their efforts. Neglecting sharing these reports can increase the feeling that data collection and reporting is an arbitrary task of little value, and this is present in the literature. Ndoye and Parker (2010) found that data usage is expanded if data is shared amongst stakeholders in "a forum in which to discuss results, practices, and findings" (p. 38).

In short, organizations that have strong leaders who prioritize data use, who plan their data use, sharing, and management activities, and who tie those plans into broader organizational goals, tend to have a more robust and rigorous data program.

Procedures

Explicit data management procedures are often lacking in nonprofit organizations, as data is often managed on an ad hoc basis based on external requirements that vary from funding source to funding source and from year to year. However, a data management strategy cannot succeed without detailed data management procedures. In fact, Mattia (2011) found that organizations that had "clearly defined and documented guidance procedures progressed to a greater overall management maturity level than those without" (p. 130).

Data management procedures can be varied and must cover a wide variety of topics. One such topic is metadata, or data about data, which describes and defines the data itself. Khatri and Brown (2010) explain that metadata plays a significant role in how data is discovered and used by organizations and, as such, must be standardized and recorded accurately, in a timely manner, and must be as complete and accurate as possible to be useful. Once defined, organizations must develop processes to ensure data is of a high quality (that it is accurate) and is secure, both essential steps in a data management plan (Levitin & Redman, 1998; Malik, 2013). Data access is also an important aspect of data management, as users need to be able to access quality data for decision making (Levitin & Redman, 1998, p. 92). And finally, an abundance of explicit procedures regarding data quality, access, security, and definitions would be useless without monitoring compliance with those procedures. As such, regular compliance evaluation should be a part of any data management plan (Thompson et al., 2015, p. 320).

Budget

While it may seem obvious that data management activities require funding to operate, there is little in the literature that describes the link between funding and successful data management. Bernard and Pukstas (2009) found that self-identified technology leaders in the 388 nonprofit sector respondents spent 2.5 times more on their information technology than others in the survey. Similarly, Carman and Fredricks (2010) found that organizations that lacked access to technology struggled with evaluation activities. Researchers also point to a lack of funding and support for data management tasks, such as storage and analysis, as being a significant barrier to good data management practices (Beath et al., 2012; Stoecker, 2007).

Botcheva, White and Huffman (2002) found that external funding was not significantly correlated with the consistency of data collection, however, this could be due to the later findings of Carman and Fredericks (2008) who found that only 8% of their respondents reported that funding for evaluation activities was included in their grant funds, and none reported receiving funds specifically for evaluation activities. This describes a situation in which evaluation, a required and data-heavy activity, is hampered by the lack of direct funding for those activities. This conclusion is supported by research from Stoecker (2007) who found that most funders do not provide enough support for organizations to conduct front-end (needs assessments, asset assessments) or back-end (program evaluation) data activities (p. 114).

In addition to a lack of funds for data management, poor data management can lead to a loss of revenue for organizations. Stoecker (2007) found that a lack of trained personnel lead to wasted time in organizations, as more time was spent finding and correcting errors when staff were untrained (p. 111). Redman's (1998) analysis of studies on costs of poor data quality found that anywhere from 8% to 12% of revenue can be consumed by data quality issues or that 40-60% of expenses can be attributed to fixing problems associated with data quality (p. 82).

The link between data management budgets and the quality of a data management program is limited. Mattia (2011) did find that data management activities that were budgeted via line item that "coordinated management of enterprise data activities" were associated with a greater level of data management maturity than those that did not budget data management. This agrees with previously held notions that as data processing growth moves to later stages, level of data processing expenditures likewise increase (R. L. Nolan, 1979). It is important to note that some studies have come to the opposite conclusion. One study tested Nolan's 1973 stage model using various county budgets from California, and found no support for the concept of increased budget along stages of data processing (Lucas & Sutton, 1977; Mattia, 2011). However, the Mattia (2011) research went further than Lucas and Sutton (1977) as it studied "management activities budgeted according to line items in a manner permitting coordinated management of enterprise data activities" (p. 130).

An organization's access to technology that enables data management is related to budgeting.Stoecker (2007) suggested that even with better funding and data collection expertise, "nonprofits are likely to still face capacity challenges in collecting and using good data" as access to usable databases are limited (p. 114). One of the issues related to technology described by Andrei, Pope, Hart and Quinn (2012) in their survey of 467 nonprofit professionals, was that data storage methods usually hamper accessibility. The authors explain that 46% of respondents described data was stored in multiple areas, and the authors suggest that software that can house and manage data in one place would alleviate some analysis and use challenges. The same study found that 42% of respondents reported no access to adequate analysis tools.

Theoretical Framework: Stage Theory

Stage theory is a useful theory to begin to understand “a pattern of specific stages that elements in systems move through over time” (Mattia, 2011, p. 123). Identification of patterns as stages is helpful as data management practices, being complex and varied operations, tend to build on previous iterations as they mature within organizations. Two key components of stage theory, as described by Nolan (1973) are that a stage theory has identified elements which can be “specified by a set of attributes,” and that it can demonstrate how those identified elements change over time (p. 400). Stage theory was first used in data management literature by Nolan (1973), who introduced a four-stage theory which described the stages of managing computer use. Their initial model included the following four stages:

- Stage 1: Initiation. Due to necessity, simple software and hardware is introduced to an organization for data processing, and personnel tend to take a hands-off approach to the new technology.
- Stage 2: Contagion. Management begins the process of explaining computer operations to personnel, and adoptions of management technology begins to take its hold in the organization. In this stage, a rapid growth of new and better technology is observed. Planning is still not recognized in this stage.

- Stage 3: Control. Because of rapid growth in technology, computer operating costs are curtailed, and project management activities are prioritized. This is usually initiated after a crisis, typically associated with costs. While computer use continues, organization, IT importance, and controls are present. Formal project management is a hallmark of this stage.
- Stage 4: Integration. Control by end-users is on the rise, data processing budget increases, and there is a greater demand for online databases as data is used more frequently. This stage is important as formal planning and control are present within data processing, representing a data management strategy.

Since Nolan's initial model, later studies have utilized stage theory to attempt to understand technological advances within organizations with mixed success. Mattia (2011) utilized a stage model to test the data management maturity of organizations that budget for data management practices and that have “defined and documented” data management procedures (p. 126). The stages of maturity for their study include:

- Stage 1: Initial. Application level data administration encouraged, data quality problems are identified and corrected, and data support is available.
- Stage 2: Repeatable. Planning drives data administration activities, procedures for data definitions exist, standardized data quality procedures are developed and deployed.
- Stage 3: Defined. Quality data services are provided, and standards for requesting data, enterprise level data support exist.
- Stage 4: Managed. Data administration management is introduced, data integration patterns are observed, data steward council manages data challenges, data support is coordinated by council.

- Stage 5: Optimizing. Data becomes complete and accurate, information flows effectively. Feedback is used to revise data processes. Evaluation of data exists, and proactive management of data needs is ongoing. Data support activities are optimized (p. 125).

Results from surveys suggest that those organizations that budget data management practices “progressed to a greater overall data management maturity level” and that organizations with defined and documented procedures were also further along the maturity model than those without (p. 130).

Farah (2017) presented a similar model which outlined the stages of maturity for big data management within organizations.

- Level 1: Initial (Pre-Contemplation). Ad hoc approach to big data management, lack of awareness that data management would be beneficial.
- Level 2: Defined (Contemplation). Emergence of a need for the organization to change. Additional data and data tools added, these tools used to explore possibilities of the organization and to inform strategic operations.
- Level 3: Managed (Preparation). Resources dedicated to changing the way the organization manages big data. Metadata systems and data integration deployed to assist organization. Personnel and monetary resources dedicated to big data management.
- Level 4: Optimized (Commitment). Higher level data management need recognized. Enterprise data management architecture present and illuminates a view of the data as it relates to the organizational goals and objectives.
- Level 5: Strategic (Future). Big data viewed as essential asset to organization. Big data used to enhance revenue. Data used for the strategic advantage of the organization.

One main difference between the model presented by Farah and that by Mattia is in stage one. While Mattia defines stage one as an initial phase which includes activities such as "applications of data administration" and "tool level support exists" (p. 125), Farah (2017) allows for a preconception phase of data management. This includes those organizations who "lack awareness or denial of the value that a big data program might contribute to the wellbeing of the organization" (p. 13). This allows for inclusion of organizations who may use big data on an ad hoc basis, but who do not actively attempt to manage, protect, or ensure quality of that data in a strategic manner.

Stage theory can provide an understanding of how data management concepts progress in an organization. Nolan (1973) explains that stage theory presumes that distinct stages change over time and are describable, and that knowledge gained through this study provides "a base for prescriptive theory formulation" (p. 400). Thus, studying the elements of a nonprofit's data management and how those elements change and mature over time could lead to guidelines, suggestions, or educational models which the organization could use to develop effective data management plans and strategies. Additionally, tying data management maturity levels to organizational outcomes could lead to a better understanding of the ways in which data management practices impact outcomes for nonprofit organizations and their clients.

Nonprofit Data Management Stage Model

The literature illuminated various data uses and challenges that nonprofits face when managing data. Categorizing the emerging data management activities into four domains, culture, leadership and planning, procedures, and budget, enables defining specific activities that each contribute to an overarching data management strategy. Since, as Nolan (1973) described, stage theory examines how elements change over time, the following model of nonprofit data management stages has been categorized into five distinct stages: (1) pre-conceptual, (2) conceptualization, (3) preparation, (3), dedicated management, and (5) strategic management. At the (1) pre-conceptual stage, nonprofits use data ad-hoc with little to no concept of how data can be used to benefit their organization beyond required reporting (such as to funders). Data practices are not assigned to specific positions; instead, data tasks are assigned to individuals and may change frequently. Little or no funding is budgeted for data activities, and any related technology funding (such as computer or software) are used for multiple functions rather than dedicated to data activities. At the (2) conceptualization stage, organizations realize that data can be used for more than required reporting. Specific individuals are assigned to data tasks and take ownership of those tasks, but formalized roles for data management are still missing from job descriptions. At this stage there is a recognized need for additional data management and data use. At the (3) preparation stage formal conversations about how to better manage data occur between key staff. Job descriptions are updated to include data management tasks. Data is regularly used for internal decision-making, and there is a formal plan to begin documenting data procedures. At the (4) dedicated management stage, there is documentation that outlines data procedures, and data is used for organizational decision-making. Additional budgeting is created for data tasks, such as technology and training required for adequate IT infrastructure. At the (5) strategic management stage, organizations use data to make decisions at the organizational and departmental levels, and spend additional time and resources to explore data for innovative solutions to problems. Strategic initiatives at this level shift from implementation to improvement and maintenance of data management systems and practices. Regular audits occur to ensure data quality and compliance with company procedures. In essence, data is seen as a valuable resource and is treated accordingly.

At each stage, data management elements fall under one of the four domains described above. Thus, specific data use and management activities can be tracked as they develop from one stage to the next.

The model proposed in this study has been inspired by the Farah (2017) model which includes a pre-contemplation stage. Literature shows that some nonprofits collect and analyze data only to satisfy external funding requirements, thereby only considering data collection and reporting an activity that must be performed in those specific cases. The later stages of the proposed model show an organization that grows to become data-driven, with internally motivated data activities, trained staff with ownership and accountability, clear procedures, dedicated funding for management activities, and a data management plan that is tied to the broader strategic goals of the organization. The complete Nonprofit Data Management Stage Model is displayed in Table 1.

Table 1

Nonprofit Data Management Stage Model

	Culture	Leadership & Planning	Procedures	Budget
Sage 1: Pre-Conceptual	Data used only as far as is required by funders or regulations. Data is viewed as useful only for securing funding. Staff are not discussing data management plans	Data management tasks are not in job descriptions and there are no plans to review or update them. No data training process in place for staff. Organization does not have a strategic plan, or if one is in place, no data management goals are present.	No formal data management procedures exist or if they do, they are not known by most staff. Data is not checked for accuracy after it is created.	Organization feels that money should only be spent on data management technology that is multi-purposes. Budget does not include items beyond hardware and software. No permanent budgeted line item for data management equipment or tasks. Adata training not included in the organizational budget.
Stage 2: Conceptualization	Data used for funders and to satisfy regulations, but the organization recognizes that its use can be expanded. Data is being used for a few other tasks (such as program planning, service delivery, budgeting, funding search/application, and supervision). Staff rarely, discuss data management tasks or plans	Data management tasks are not in job descriptions, but talks are occurring around reviewing and updating. Data training not available beyond on-the-job training for project-specific tasks. Organization has a strategic plan and there are goals to start data collection and/or use.	Formal data management procedures exist, but only because regulations require it. Data management procedures are in place for some data and include at least one of the following: protection against unauthorized access, data accuracy, data backups, data analysis, data sharing and reporting, and metadata. Data backup procedures are in place for some data, but procedures are not complete and only contain one of the following: how often to backup, where to backup, and how to backup. Data is only checked for accuracy if a problem is discovered.	Organization feels that money should be spent on data management technology only when required for a specific project or task. Budgets include funding for hardware and software and are starting to include expansion of those items. Budgeted training is limited to only those who use data regularly. Data training for staff is not in the budget, but on-the-job training provided by staff is available for those who use data regularly.
Stage 3: Preparation	Data used for funders and to satisfy regulations, but the organization is talking about or making plans to expand its use. Data is increasingly used for other tasks (such as	Data management tasks are not in job descriptions, but plans are being made to update or updates are in process. On-the-job training available for program	Data backup procedures exist for some data and where it exists is complete, containing all of the following: how often to backup, where to backup, and how to backup. Data management	Organization feels that money should be spent on data management technology whenever funding is available. Budget funds some aspects of data management but is limited to training (though

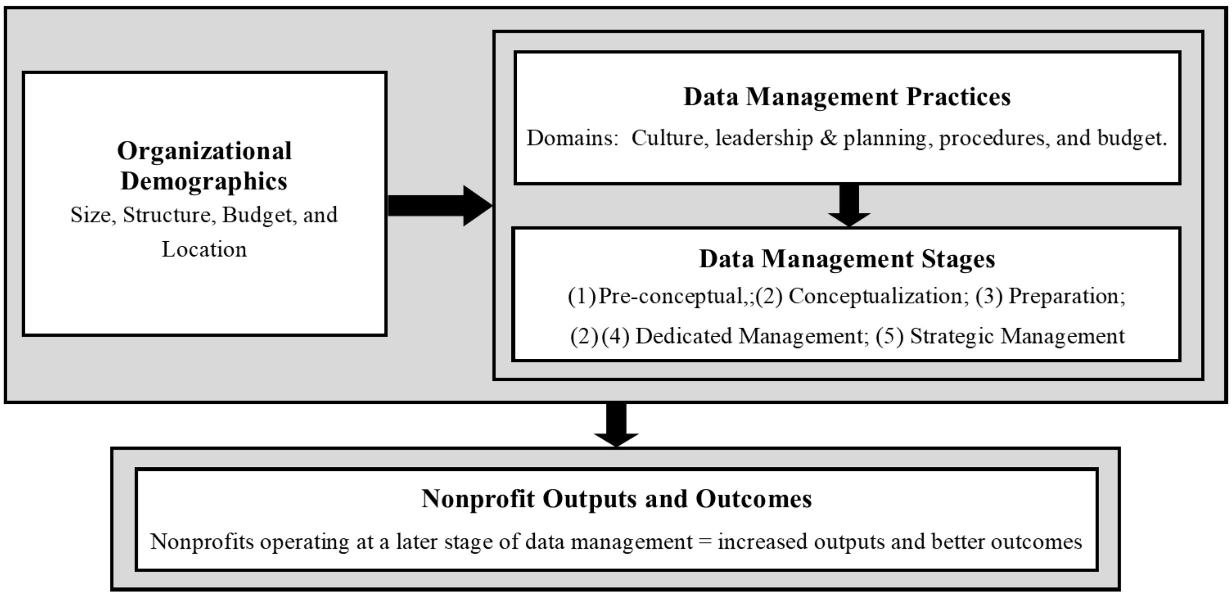
	program planning, service delivery, budgeting, funding search/application, and supervision). Staff spend some time discussing and planning to improve and innovate data management.	specific and non-program specific tasks. Organization has a strategic plan that includes goals to improve data collection and use.	procedures are in place for some data and include at least two of the following: protection against unauthorized access, data accuracy, data backups, data analysis, data sharing and reporting, and metadata. Data is checked for accuracy occasionally or if a problem is discovered.	now open to all staff), software, and hardware, as opposed to money for tasks such as evaluation or data quality audits. Data training is not in the organization's budget, but on-the-job and paid training is available for staff.
Stage 4: Dedicated Management	Data used for more than just funders and to satisfy regulations and the organization is carrying out plans to expand its use. Data is used regularly for other tasks (such as program planning, service delivery, budgeting, funding search/application, and supervision). Staff frequently spend time discussing and planning to improve and innovate data management.	Data management tasks exist in job descriptions for most employees who manage data. Paid training available for program-specific tasks. Organization has a strategic plan that includes goals to expand use of data beyond current practices.	Data management procedures are in place for some data and include most of the following: protection against unauthorized access, data accuracy, data backups, data analysis, data sharing and reporting, and metadata. Data backup procedures exist for all data and are close to complete, containing one or two of the following: how often to backup, where to backup, and how to backup. Data is checked for accuracy on an occasional basis, but not regularly.	Organization feels that there should be a budget for required data management technology. Budgets include money for many aspects of data management, including IT software and hardware expansion, program evaluation, and consulting. Data quality audits are not a part of the agency budget at this stage. Data training is budgeted for staff who use data regularly.
Stage 5: Strategic Management	Data used in most aspects of operations and the organization is consistently working to make data collection and use more efficient. Data is used for all aspects of the program, including program planning, service delivery, budgeting, funding searches/application, and supervision. Staff spend time discussing and planning to improve and innovate data management.	Data management tasks exist in job descriptions for all employees who manage data. Paid training available for program specific and non-program specific tasks. Organization has a strategic plan that includes goals to explore innovative ways to use data.	Data management procedures are in place for all data and include most or all of the following: protection against unauthorized access, data accuracy, data backups, data analysis, data sharing and reporting, and metadata. Data backup procedures are in place for all data and contain all of the following: how often to backup, where to backup, and how to backup. Data is checked for accuracy on a regularly scheduled basis.	Organization feels that there should be a budget for both required data management technology as well as new or innovative data management technology. There is a permanent organization-wide budget line-item for data management tasks in place. Budgets include money for most aspects of data management, including IT software and hardware expansion, program evaluation, data quality audits, and consulting services. Data training budgeted for both data users and staff interested in data.

Conceptual Model

The conceptual model that guides this research shows the relationship between data management practices and identified outputs and outcome measures, as well organizational demographic influences including budget, size, structure, and location. The conceptual model can be seen in Figure 1 below.

Figure 1

Nonprofit Data Management Conceptual Model



The conceptual model presented above shows the relationship between the data management practices of an organization and their place in the nonprofit data management stage model. These data management practices include four data management domains which are described in detail below. The data management practices within the four domains were placed within one of the five data management stages which are further defined below. According to this model, where an organization is operating within the data management stage model will impact their outputs and outcomes. For example, a nonprofit operating at later data management stages should see increased outputs and better outcomes than those operating at the earlier stages.

However, data management practices do not exist within an organizational vacuum. Organizational factors such as size, (fewer than 10 employees vs. large with hundreds), structure (traditional top-down management vs. circular management), and budget (large vs. small), all impact the practical implementation of a data management plan.

Conclusion

This chapter introduced the Nonprofit Data Management Stage Model, informed by the literature review, which guided the survey development process. The model included four data management domains, culture, leadership and planning, procedures, and budget, and listed specific elements that are in each domain that change over time as a nonprofit moves from stage one to stage five. Finally, the conceptual model that guided this research was presented, outlining the relationship between the nonprofit data management stage model and nonprofit outcomes and outputs while considering other factors such as organizational demographics.

Chapter 3: Methodology

Introduction

Chapter 3 outlines the research design for this dissertation including the research question and resulting three hypotheses. In addition, this chapter presents the study population, the process for selecting subjects, the survey instrument, and a description of how participant responses were scored for later analysis.

Research Design

Due to limited previous research on this topic, this study was exploratory in nature. The literature review was used to illuminate the potential data collection activities and uses within nonprofit organizations. General conclusions from the literature were organized into themes, which were used to develop a Nonprofit Data Management Stage model that identified the various data management activities that nonprofits undertake in five distinct stages. Next, a series of survey questions were developed to explore the data management activities of the survey sample, which was domestic violence (DV) organizations providing emergency shelter services to DV survivors. This population included all DV shelters within the United States. The survey also requested that respondents provide basic output and outcome data for their organization as well as demographic data from their organization, the individual respondent, and the populations they serve. This survey data was used to explore the following:

1) The overall accuracy and quality of the Nonprofit Data Management Stage Model.
2) The relationship between agency outputs and outcomes and their data management maturity level (based on the stage that they were placed into).
3) The relationship between the organizational demographic factors identified in the conceptual model and the agencies' data management maturity level.

4) The relationship between an organization's data management maturity and how often it uses data for decision making.

While DV shelters offer a uniform and comparable sample, outcomes tracked by DV shelters are minimal. Sullivan (2012), in a review of outcome evidence for DV shelters, found that while shelters were found to be supportive and effective, the evidence used was mainly self-reported data that was not uniformly collected from all shelters. In fact, discovering regularly reported outcomes for DV shelters proved to be a significant challenge. One of the main funders for DV shelters in America is the US Department of Health & Human Services via the Family Violence Prevention and Services Program (FVPSA) grant. This grant requires uniform reporting of outputs, such as the number of individuals served, demographic data of clients served, the number of services provided, and the number of presentations provided to the community, but the outcomes required are less defined and rely, again, on self-reported surveys (US Department of Health and Human Services [USDHHS], 2012). Sullivan (2011) suggests that simplified outcomes applied to all DV clients would make outcome measurement activities more attainable for DV service providers. The basic outcomes suggested are "(1) survivors will increase their knowledge about community resources available to them, and (2) survivors will have strategies for enhancing their safety" (p. 356). While these suggestions may be simplified and more easily collected and reported, they are not currently required to be collected by DV shelters, and therefore, were likely unavailable for this study. Since uniform outcomes were not identifiable in the research or via funder requirements, a combination of output data and outcomes developed through careful consideration of what was likely regularly collected was used.

The survey collected the following outputs and outcomes that were used in analysis as the dependent variables.

Selected outputs included in this study were those identified by the Family Violence Prevention and Services Administration, the governmental organization that provides funding for DV shelters across the nation. Those outputs include the number of individuals served who resided at the shelter, the number of individuals served who did not reside at the shelter, the total number of legal advocacy services provided to individuals, and total number of community presentations provided.

Outcome data requested was the average length of time a client spent in the shelter and the percent of clients that exited the shelter to stable housing.

Independent variables include agency demographics, including budget, size, location, and structure, and the stage model scores assigned to each respondent based on their responses to the survey instrument. The stage model scores were also used as dependent variables when the relationship between agency demographics, such as budget and size, and the stage model score was explored. Finally, a revised 2-factor stage model was used after a psychometric analysis was performed on the collected data.

Research Questions and Hypothesis

This dissertation attempts to answer the following question: are data management practices related to nonprofit program outcomes? To answer this question, three hypotheses were tested:

- Hypothesis 1: Nonprofits at a later stage of the nonprofit data management stage model will have increased outputs compared with those at an earlier stage.
- Hypothesis 2: Nonprofits at a later stage of the nonprofit data management stage model will use data for decision making more than those at an earlier stage.

- Hypothesis 3: Nonprofits at a later stage of the nonprofit data management stage model will have better program outcomes than those at an earlier stage.

Study Population and Sample

Domestic violence (DV) housing service programs are in operation in every state in the United States. According to domesticshelters.org (2018), a nonprofit organization that complies and verifies information about domestic violence shelter and other service providers in the United States and Canada, 2,459 programs in the US provided DV housing services at the time of this study. Every state in the US has a DV coalition which provides information and resources for individuals experiencing domestic abuse, as well as support to those agencies providing services.

DV housing programs are an appropriate study population for multiple reasons. First, DV funding has been on the rise in recent years due to the passage of the Violence Against Women Act in 2000, which was reauthorized in 2005 with the addition of housing provisions such as the prohibition of evictions based on perceived DV, or the inclusion of housing voucher portability in cases where a move is for safety reasons (Baker, C. K., Billhardt, K. A., Warren, J., Rollins, C., & Glass, N. E., 2010). This stabilized funding coupled with (and likely a component of) the existence of DV housing programs in every state, make the study population consistent. Second, while each organization's demographics will have unique partnerships, programs, practice models, and populations, all provide similar services with similar outcome goals (providing housing for DV survivors) thereby making comparisons between organizations more feasible. Finally, shelters typically serve clients with similar problems and similar dynamics, making comparisons less problematic. For example, clients seeking DV shelter typically seek services because they need safe and stable housing, legal services, counseling, and support services as well as case management.

Recruitment of Subjects

Due to lack of research in this area, no existing survey collects information related to nonprofit data management practices. As such, a survey was developed to discover the data management practices of selected nonprofits. The survey was delivered via the online survey platform Qualtrics to a list of DV shelters across the nation. Two survey contact lists were created. The first was populated by obtaining a list of DV coalitions from each state from the National Coalition Against Domestic Violence website. Each DV coalition website contains a list of DV shelters within their state. From these lists, shelter websites were searched to locate email addresses of, preferably, the executive director of the DV shelter. If the executive director email was not published, the general shelter email address was used. There were 691 DV shelters with adequate contact information used in this study. The second list was populated by finding the contact information for the 48 statewide domestic violence coalitions from the United States. This list was used to request that coalitions send the survey link to shelters in order to raise the likelihood of shelters responding, and to expand the list of shelters reached.

The recruitment schedule is outlined below:

1. List of DV shelter contacts created.
2. List of DV statewide collation contacts created.
3. Recruitment email sent to DV shelter contacts that describes the research and requests that they reply with the most appropriate email address to send the survey to.
4. Each DV statewide coalition was called and the researcher described the research and asked the coalition to send out the survey link to their shelter email list. An email

with the survey link for their state was sent once the coalition agreed to review and send the survey.

5. Qualtrics links were created for each state. An email contact list was populated within Qualtrics for each DV shelter within a state, and reminder emails were sent out from the Qualtrics system.
6. An anonymous weblink was provided to each coalition to send out to their respective email lists.
7. Respondents were anonymous, and no identifying information was saved.

Response Rate and Missing Data

During the recruitment of subjects, five coalitions did not respond to emails or phone calls, and two expressed that they did not want to be a part of the research. The remaining states asked to be sent the survey and indicated that they would consider sending it out to their email lists. 43 state coalitions were ultimately sent the survey link and instructions. Of those 43 coalitions, 17 states sent out the survey to their email lists totaling an estimated 572 shelters.

The second survey population was the domestic violence shelters themselves located in the United States. The list of every domestic violence shelter in each state that was created using coalition websites was used, and a total of 1,590 shelters were located. Of the 1,590 shelters, a total of 592 emails from 46 states were publicly available. Four states did not list contact information for staff at their shelters or for the shelter itself. The remaining states had some shelters that listed staff contact information, and others that provided only an organization-wide general email address. Of the 592 emails that were sent out during the study, 27 emails bounced or otherwise failed to send, and 29 respondents opted out of emails from Qualtrics. The final number of surveys successfully sent was 536.

Since some shelters likely received the survey two ways, from their statewide DV coalition and from a direct email from Qualtrics, the total number of individual emails sent out was estimated to calculate the survey population size. To estimate the total number of shelters who received the survey via email, we took the total number of surveys sent to individual shelters (536) and added the total number shelters in the states whose state coalition sent out the survey link that did not have published emails online (342). The total survey population estimate was calculated to be 878.

While this research used a variety of analysis methods, it is important to remember that the response rate, which is discussed in the next chapter, was lower than desired for most analysis methods. The final number of respondents included in the study was 73.

Of the 73 respondents who completed the survey, not all completed every question. Multiple methods were employed to manage missing data in these cases. For the survey questions that required respondents to give a number, such as the number of clients served, the series mean was used in place of missing data. The method was used in eight or less cases for each of the 11 questions that had missing data. For most questions, substitutions represent less than 10% of the total cases. Table 2 shows the number of respondents for each of the 11 questions in which series mean was used as a substitute for missing data.

Table 2

Series Mean Occurrences

Question	# of missing responses	Percent of substitutions
Budget dollar amount	4	5.48%
% budget from gov. grants	5	6.85%
% budget from private grants	6	8.22%
% budget from donations	6	8.22%

% budget that is flexible	8	10.96%
# clients served in shelter	3	4.11%
# of clients served out of shelter	1	1.37%
# of legal advocacy services provided	6	8.22%
# of community presentations	3	4.11%
Average length of state in shelter	3	4.11%
% of clients exit to stable housing	5	6.85%

For questions that asked the respondents to indicate if their organization performed tasks or action, a zero was put in place of missing data. The reasoning is that none of the 73 respondents included in the study failed to finish the survey. Because of this, it can be safely assumed that if the respondent answered some of the process questions, the questions that were skipped were either not performed by the organization or that the respondent did not understand the question, indicating it is not a topic that is discussed in their organization and therefore likely not an action that is performed regularly, if at all.

Instrumentation

Survey

The survey instrument consisted of 50 questions with 132 separate items using a combination of open answer, multiple choice (single option), multiple choice (multiple selection), and matrix questions. Survey questions explored data management practices within the four identified practice domains including organizational culture, leadership and planning, budget, and data management procedures. Additional questions collected information about the organization itself, the data the organization collects, and the survey taker. Data collected from the survey regarding activities occurring within the practice domains helped to place organizations within one of the five data management stages. Since there is not an existing nonprofit data management survey in existence, additional factor analysis was conducted on the survey responses to verify the construct validity of the survey instrument. The final survey instrument can be found in Appendix A.

Data Collection

The survey was open to respondents from October 31, 2018 through January 15, 2019. Data was downloaded from the Qualtrics survey platform and cleaned using SPSS 24 statistical software. The complete dataset was also converted to an Excel file where most of the data cleaning and scoring calculations were performed prior to data analysis.

Data Analysis

Analysis of the data for this research included a psychometric analysis of the survey instrument as well as an analysis of the data collected to explore the relationship between data management practices and program outputs and outcomes.

Psychometric Analysis

A scoring mechanism was developed using the survey instrument and the collected data. Scoring included a score for each of the four practice domains, as well as an overall score used to place respondents into one of the five data management stages. A detailed description of the scoring plan is outlined in the following section.

Once scoring was complete, an Exploratory Factor Analysis (EFA) was conducted using a principle components analysis with varimax rotation. Since the instrument used for this study is new, and the research that guided its development was incomplete as nonprofit data management practices have not been thoroughly studied, the EFA was used to explore the factorial structure of the instrument itself.

After the EFA was complete, the rotated component matrix was used to remove items that did not load into one of two primary factors. Once the items were identified, a new scoring protocol was developed, and new factor scores were calculated. An additional reliability analysis was performed using Cronbach's alpha reliability scale.

Descriptive Statistics

Organizational demographics, including size, structure, budget, budget funding sources, leadership structure, and client populations served, were analyzed using frequency and descriptive analysis via SPSS. In addition, frequency and descriptive analysis was performed on individual respondent demographics, including the number of years they have worked with their organization, their educational attainment, and a list of the formal data training they had received prior to taking the survey.

Additional descriptive analysis was performed for organizational outputs, including the number of clients served in the shelter, the number of clients served not residing in the shelter, the number of clients who received legal advocacy services, and the number of community presentation provided by the organization, as well as organizational outcomes, including the average length of a client's stay in the shelter and the percent of clients who exit the shelter to stable housing.

Regression Analysis – Hypothesis One and Three

A regression analysis was performed to test hypotheses one and three,. A multiple linear regression was conducted to explore the relationship between the revised 2-factor stage model for an organization and their program outcomes (length of stay in shelter and percent of clients that exit the shelter to stable housing) as well as their program outputs (total clients served in-shelter and out of shelter, total legal advocacy services provided, and number of community presentations given). This was performed using the revised 2-factor stage model as the independent variable (lower scores = less robust data management practices while higher scores = more robust practices). Two additional variables were included in the model: organization size (number of paid employees) and organization budget, as additional independent variables. Since the total number of survey respondents was too small to confidently rely on the factor analysis results, the original stage model total score was also used to explore the relationship between the stage model total score and the dependent variables described above.

Prior to running the regression analysis, survey items were checked for skewness and kurtosis using SPSS. Based on the results, a logarithmic transformation was performed to overcome moderate and substantial skewness in the data. Tables 3-5 below show the results of the logarithmic transformations. Each variable was tested using a formula for both moderate and substantial skewness, in order to find the most reliable formula for each variable. The highlights show the model ultimately selected for the regression analysis.

Table 3

Logarithmic Transformations of Program Outputs

Program Outputs	Raw		Moderate		Substantial	
	Skew	Kurtosis	Skew	Kurtosis	Skew	Kurtosis
# Reside at Shelter	1.201	1.000	.277	-.565	-1.540	3.279
# Not Residing at Shelter	4.037	18.261	2.310	6.366	-.814	4.151
# Legal Advocacy Services	4.982	28.963	2.497	8.356	-.861	.723
# Community Presentations	4.550	24.824	2.338	7.109	-.346	.874

Table 4

Logarithmic Transformations of Program Outcomes

	Raw		Moderate		Substantial	
	Skew	Kurtosis	Skew	Kurtosis	Skew	Kurtosis
Avg Length of Stay in shelter	3.891	20.806	1.249	5.250	-1.674	6.273
% Exit to Stable Housing	-.679	-.439	.557	-.581	.435	-.693

Table 5

Logarithmic Transformations of Independent Variable

	Raw		Moderate		Substantial	
	Skew	Kurtosis	Skew	Kurtosis	Skew	Kurtosis
Budget Dollar Amount	3.572	14.014	1.910	5.389	-.158	.858

Spearman's Rho - Hypothesis Two

For hypothesis two, which states that nonprofits at a later stage of data management maturity will use data for decision making more than those at an earlier stage, a Spearman's Rho test was performed. The stage model total score and revised 2-factor model were each used to test the association between data management maturity and the use of data for decision making.

Scoring – Original Stage Model

Each respondent answered a number of questions that related to various elements in the stage model. Scoring was done first by domain (culture, leadership and planning, procedures, and budget) and then by taking a total score for all domains combined to give a stage model total score.

Each domain involved multiple survey questions, some of which involved matrix options, requiring a variety of scoring approaches. A total of 14 survey questions using 55 items were used to score participants. While scoring protocol varied slightly based on the type of survey question, each domain's total score was used to categorize their stage in the model. After calculating the stage model total score, a protocol was used to develop a range of scores that assign a respondent to a stage.

To calculate the stage model total score, each score for the survey questions related to the given domain were summed to get a total score. Next, a scoring protocol was developed to assign respondents to a stage in the model. A lower score represented less mature data management practices, and a higher score more mature practices. To determine where an organization fit in the stage model, a range of possible scores was assigned to each of the five stages. Since each question score represents a stage in the stage model, summing all three scores provided the minimum possible score a respondent would need to be sorted into the corresponding stage. For example, in order for a domain with three survey questions to be assigned to stage one in the model, a respondent would have to score one on all three of the survey questions, so the minimum total score they could receive would be three (Q1[1] + Q2[1] + Q3[1] = 3). The minimum total scores a respondent could receive to be placed into stage two in the example domain is six, since responses corresponding to stage two in each of the three questions received a score of two and summing the scores for each response equals six (Q1[2] + Q2[2] + Q3[2] = 6). By using this minimum score as a starting point, a range for each stage was created, encompassing the minimum score for that stage up to the minimum for the subsequent stage. So, for stage one, a respondent could have a total score of three (the minimum score for that stage), four, or five (the maximum score before reaching the minimum score for the next stage). For stage two, a respondent could have a score of six, seven, or eight. For stage three, a respondent could have a score of nine, 10, or 11. For stage four, a respondent could have a score of 12, 13, or 14. The maximum score of 15 would place an organization into the fifth and most advanced data management stage.

Table 6, below, provides a summary of the 14 survey questions and the scoring protocol. A detailed description of each domain, the survey questions used, and the scoring protocol for those questions follows.

Table 6

Domain Scoring Protocol

Domain	# of Survey Questions	# of Items	Score range for each stage				
			Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
Culture	3	14	3-5	6-8	9-11	12-14	15
Leadership & Planning	3	14	3-5	6-8	9-11	12-14	15
Procedures	4 (two skip-logic q’s)	17	4-7	8-11	12-15	16-19	20
Budget	2	10	2-3	4-5	6-7	8-9	10

Culture

Data management practices, budget decisions, and policies and procedures are intrinsically linked to the culture of the organization. An organization with a culture that emphasizes data management, data use, and data collection will have more robust conversations, planning sessions, and buy-in for advanced data management practices. The Nonprofit Data Management Stage Model identifies several data management elements that fit into the culture domain, including the amount of time staff spend talking about and planning data management tasks and improvement, the perception of how useful data is as it relates to practices like service delivery and program planning, and how data is actually used in day-to-day operations. To measure each organization's culture domain stage, the survey asked three questions tied to the elements within the stage model. The first was a matrix-style question which asked respondents to select whether they never, rarely, sometimes, frequently, or almost always performed each of the tasks listed below:

- Our staff spends time talking about how to use data to meet reporting requirements.
- Our staff spends time talking about the need to use data more effectively.
- Our staff spends time making plans on how to use data more effectively.
- Our staff spends time carrying out plans on how to use data more effectively.
- Our staff spends time in continuous quality improvement about how to manage data better.
- Our staff spends time using data to answer questions about our programs and services.
- Our staff spends time exploring data to find patterns or information that will help us improve our services.
- Our staff spends time researching new or innovative ways to manage data.

Culture – Question One Scoring. Each of the statements above represent data management tasks that would be performed regularly in organizations with a culture that supports and emphasizes data management and use. As such, scoring for this question was accomplished by summing the total number of “never,” “rarely,” “sometimes,” “frequently,” and “almost always” responses that a respondent selected for each of the statements. Respondents who selected “never” more frequently than any response were given a score of one, which corresponds to stage one in the stage model. Respondents who selected “rarely” more frequently than other responses were given a score of two, and so on for each selection. Respondents who scored equally in two options were sorted into the higher of the two stages (so if a respondent had three responses each in both of the “sometimes” and “frequently” categories, they were given a score of four, which corresponds to the “frequently” selection).

The next question asked respondents how their organization as a whole viewed data. Respondents could select as many of the six options as they felt was appropriate for their organization. The options corresponding directly to elements within the Nonprofit Data Management Stage Model under the culture domain, as follows:

- Data is viewed as useful for getting funding.
- Data is viewed as useful for service delivery.
- Data is viewed as useful for program planning.
- Data is viewed as useful for budget decisions.
- Data is viewed as useful for supervision.

Each statement represents a use for data, and the more statements an organization selected, the more advanced their understanding of data usefulness is presumed to be.

Culture – Question 2 Scoring. For the question above, respondents were given higher scores for selecting a greater number of options. If only one option was selected, respondents were given a score of one, which corresponds to stage one in the stage model. If two options were selected, a score of two was given, if three options were selected, a score of three was given, and so on. The maximum score for this question was five.

The final question related to the culture of data management practices in the organizations asked about how organizations are currently using data. Organizations were able to select one option, and each option represented increasingly sophisticated uses of data. In addition, each response relates directly to elements listed in the Nonprofit Data Management Stage Model under the culture domain. Stage four is represented twice in the statements, as both elements are present in the stage model under stage four but were separated in the survey to avoid a double-barreled option. The statements in the survey are listed below with the corresponding stage in parentheses:

- Only as far as it is required by funders or regulations (stage one)
- Use data for funders or regulations but recognize that its use can be expanded (stage two)
- Use data for funders or regulations but are talking about ways that its use can be expanded (stage three)
- Use data for more than just funding requirements and are making plans to expand its use (stage four)
- Use data on a regular basis and are currently carrying out plans to expand its use (stage four)
- Use data in most aspects of operations and consistently works to make data collection and use more efficient (stage five)

Culture – Question 3 Scoring. Scoring for this question was simple as each statement related directly to elements within the stage model. The statement which indicates that the respondent's organization uses data only as far as it is required by funder or regulations corresponds to stage one of the Nonprofit Data Management Stage Model and was given a score of one. If the second statement was selected, the respondent was given a score of two. If the third statement was selected, the respondent was given a score of three. If the fourth or fifth statements were selected the respondent was given a score of four. And if the final statement was selected the respondent was given a score of five.

Culture Total Score. To place each respondent into a stage in the culture domain of the stage model, each of the scores for the three culture survey questions were summed to get a total score. The domain total score protocol outlined in the introduction of this section was used to calculate the score ranges for each domain, and to place respondents into a stage in the culture domain. The score ranges for the culture domain are summarized in Table 7 below.

Table 7

Culture Domain Score Range

	Score Range
Stage 1	3-5
Stage 2	6-8
Stage 3	9-11
Stage 4	12-14
Stage 5	15

Leadership and Planning

Organizational leaders that are dedicated to data management are an essential element of a data-driven organization. Leadership, however, can take many forms. The Nonprofit Data Management Stage Model suggests that leaders guide an organization through planning activities, such as strategic planning, as well as formalize important data management tasks and skills through detailed job descriptions or data management training for staff. Three survey questions were used to measure leadership and planning maturity in the responding organizations. The first question asked respondents to select a statement that best describes their organization's approach to including data management tasks in formal job descriptions. Respondents selected from the following options:

- Data management tasks are not in job descriptions and we are not reviewing or updating

- We have talked about reviewing job descriptions to include data management, but we haven't made plans to do so yet
- We have made plans or are implementing plans to review data management tasks in job descriptions
- Our jobs descriptions already include data management tasks for most employees who manage data
- Data management tasks are already in job descriptions for all employees who manage data

Leadership and Planning – Question 1 Scoring. Each of the options above correspond directly to elements in each of the five stages, in ascending order. Respondents who selected the first response were given a score of one, which corresponds to stage one. Respondents who selected the second option were given a score of two, which corresponds to the second stage, and so on.

The second question used to measure data management leadership and planning maturity asked respondents to identify the types of data training available to staff in their organization. Respondents were able to select as many of the following options as applied:

- Paid training or workshops that are program specific.
- Paid training or workshops that are not program specific.
- On-the-job training that is program specific.
- On-the-job training that is not program specific.
- No data training process in place.

Leadership and Planning – Question 2 Scoring. Scoring for this question, again is related directly to the stage model elements in each stage. If a respondent selected the last option, “no data training process in place,” then they were automatically coded into stage one and given a score of one. If the response option “on-the-job training that is program specific,” was selected, respondents were coded into stage two and given a score of two. If response option “on-the-job training that is not program specific” was selected, respondents were coded into stage three and given a score of 3 (respondents could also select any of the options from stage two). If the response option “paid training or workshops that are program specific” was selected, respondents were sorted into stage four and given a score of four (respondents could also select any options from stage two and three). And finally, if respondents selected “paid training or workshops that are not program specific” they were sorted into stage five and given a score of five (respondents could also select all other response options excluding that from stage one).

The final question for the leadership and planning domain asked respondents to select the elements that are currently in the strategic plan for their organization. Respondents were able to select as many of the following elements as applied:

- Goals to start data collection.
- Goals to start using data for decision making.
- Goals to improve data collection.
- Goals to improve data use for decision making.
- Goals to expand the use of data beyond current practice.
- Goals to explore new and innovative ways to use data.
- There are no data goals in my organizations strategic plan.
- My organization does not have a strategic plan.

Leadership and Planning – Question 3 Scoring. Scoring for this question relate directly to the stage model elements. For stage one, respondents had to select either “my organization does not have a strategic plan" or "there are no data goals in my organizations strategic plan" and were given a score of one. For stage two, respondents had to select the first or second response options only ("goals to start data collection" and "goals to start using data for decision making") and were given a score of two. For stage three, respondents could select the same response as stage two, but also had to select "goals to improve data collection" and/or "goals to improve data use for decision making" response options. If this was the case, they were given a score of three. For stage four, respondents could select any of the options for stage two or three, in addition to "goals to expand the use of data beyond current practices" response option. If they did so, they were given a score of four. For stage five, respondents were required to select "goals to explore new and innovative ways to use data" as well as any of the options from stages two to four and they were given a score of five.

Leadership and Planning Total Score. To place each respondent into a stage in the leadership and planning domain, all three survey questions were scored, and the total score was summed. Once that was done, the domain total score protocol outlined in the introduction to this section was used to calculate the ranges for each domain, and to place respondents into a stage in the leadership and planning domain. The score ranges for the leadership and planning domain are summarized in Table 8 below.

Table 8

Leadership and Planning Domain Score Range

	Score Range
Stage 1	3-5
Stage 2	6-8
Stage 3	9-11
Stage 4	12-14
Stage 5	15

Procedures

Data management procedures that are formalized and known throughout the organization are essential for data management to be consistent and reliable. For the Nonprofit Data Management Stage Model, organizations that have more mature data management practices included those that have data management and use policies and procedures, that work to ensure data collected and reported is accurate, and that ensure data is backed up and accessible. Six survey questions were used to measure the maturity of an organization's data management procedures, though two of those questions utilized skip-logic to reveal additional questions. In those cases, respondents who selected an option that led them to skip the subsequent related question were automatically scored into stage one of the stage model. If they selected an option that revealed the subsequent question, additional scoring was required.

The first question in this domain asked respondents if their organization had any policies or procedures related to data management. Respondents were able to select that "yes" they had, "no" they didn't have, or "don't know" if they had, policies or procedures for each of the following:

- Data management.

- Data sharing.
- Data use.

Once respondents answered this question, skip logic determined if they would see the subsequent question. If respondents selected “no” or “don’t know” for all of the options above, they were given a score of one, which corresponds to stage one on the stage model. If the respondents were able to view the subsequent question, they were asked to identify their organization’s motivation behind writing data management policies. Respondents could select as many of the following options as applied to their organization:

- Because regulations or laws require it.
- Our organization values data management and use.
- Our organization wanted to have a document for new staff training.
- Our organization wanted to ensure data we collect is done so the same year after year.
- Other.

Leadership and Planning – Question 1 & 2 Scoring. Respondents who were given the option to answer this question were scored by directly linking the selected response options with the elements in the stage model. Each of these responses represent motivations that correspond to a level of data management maturity within an organization. For example, organizations that go through the process of writing policies or procedures only because regulations require them to do so are seen as less invested in the data management process and therefore less advanced than those that are motivated by a desire to ensure data collection is accurate or that want to use those procedures to train new staff. Respondents who selected only “because regulations or laws require it” were given a score of two, which corresponds to stage two on the stage model. Respondents who selected “because regulations or laws require it” and at least one other option were given a score of three, which corresponds to stage three on the stage model. Respondents who selected “because regulations or laws require it” and at least two other options were given a score of four, which corresponds to stage four on the stage model. Finally, respondents who selected “because regulations or laws require it” and all other options were given a score of five, which corresponds to stage five on the stage model.

The next question in this domain related to written data management procedures that exist within an organization. Respondents could indicate that they did not have, had for some data, or had for all data, procedures for each of the following elements:

- Protection against hacking or unauthorized access.
- What to do in the case of hacking or unauthorized access.
- Ensuring data collected is accurate.
- Ensuring data is backed up.
- Describing the process by which we analyze data.
- Describing guidelines for how we report data.
- Defining data sharing requirements.
- How to define data (for example, a written description of a piece of information that is collected such as a count of unduplicated individuals to receive a service).

Leadership and Planning – Question 3 Scoring. This question scoring was based on the idea that written procedures for these data management elements are more important to an organization with more mature data management practices. The more written procedures that exist, the more mature the organization is in terms of data management. To score respondents, the number of “for some data” and “for all data” responses selected by the respondent were summed. Any respondents who selected “does not have” for either “ensuring data collected is accurate" or "ensuring data is backed up" were automatically placed into stage one of the stage model and given a score of one. This was done because these two tasks are seen as essential for any data management plan. Data accuracy and backing up data to prevent loss are foundational activities that ensure data is both available and useful. For the remaining options, respondents who selected at least two categories, but no more than six categories, and indicated that they had procedures for those categories for some data collected, were placed into stage two, and given a score of two. Respondents who selected at least three categories, but no more than seven categories, and indicated that they had procedures for both some data and all data in selected categories, were placed into stage three and given a score of three. Respondents who selected at least seven categories and indicated that they had procedures for those categories for some data, or for all data in no more than six categories, were placed into stage four and given a score of four. Respondents who selected at least six categories and indicated that they had procedures for those categories for all data collected, were placed into stage five, and given a score of five.

Next, we asked respondents how their organization, as a whole, worked to ensure that the data they collected and reported was accurate. Respondents could select one of the options below:

- Data is not checked for accuracy.
- Data is checked for accuracy when a problem is found.
- Data is checked for accuracy on an occasional basis.
- Data is checked for accuracy on a regularly scheduled basis.

Leadership and Planning – Question 4 & 5 Scoring. Scoring respondents in this question was slightly more complex, as data accuracy was also explored in the previous question where we asked respondents if they had a procedure for ensuring data is accurate. As such, this question was used to aid in more accurately measuring the data accuracy maturity for the organization. If a respondent indicated that they did not have a data accuracy procedure and they selected “data is not checked for accuracy” in this question, they were given a score of one, which corresponds to stage one of the stage model. This combination of not having a data accuracy procedure in place and not checking data for accuracy, indicates a lower level of data management maturity for the organization. If the respondent indicated that they did not have a data accuracy procedure, or that they did for some data only, and that they either do not check data for accuracy or that they check data for accuracy when a problem is found, they were given a score of two which corresponds to stage two on the stage model. This score implies that they organization either checks data for accuracy when a problem is found, or that they at least have a procedure in place for ensuring data accuracy. To be given a score of three, which corresponds to stage three on the stage model, an organization had to indicate that they had a procedure for ensuring data is accurate for some or all data, and had to indicate that they checked data for accuracy once a problem was found. To be given a score of four, which corresponds to stage four on the data management stage model, respondents had to indicate that they had a procedure for data accuracy and that they checked data for accuracy on an occasional basis, not just when a problem is found. And in order to be given the maximum score of five, which corresponds to stage five on the stage model, respondents had to indicate that data is checked for accuracy on a regularly scheduled basis.

The final question that measured data management procedure maturity asked the respondent to describe their data backup procedure if one exists. Respondents who indicated on an early survey question that they did not have or that they were unsure about an existing procedure for data backups were automatically given a score of one which corresponds to stage one on the stage model, and were not shown the subsequent question. For those that were showing the subsequent question, they were asked to select which elements were in their data backup procedure, and were able to select all options that applied to their organization from the list below:

- How often data should be backed up.
- Where data should be stored.
- How to perform data backups.

Leadership and Planning – Question 6 Scoring. As with the previous question, responses from two questions were combined to accurately measure the data backup maturity level of the organization. To get a score of one, as described above, a respondent had to indicate that did not have or were unsure if they had a procedure for ensuring data is backed up. In order to receive a score of two, which corresponds to stage two of the stage model, a respondent had to indicate that they had a data backup procedure in place for some data, and that they had at least one of the above elements in that procedure. In order to receive a score of three, which corresponds to stage 3 on the stage model, the respondent had to indicate that the organization had a backup procedure for some data, and that the procedure included more than two of the elements above. In order to receive a score of four, which corresponds to stage four on the stage model, respondents had to indicate that a data backup procedure was in place for all data, and that one to two of the elements above were in that procedure. And finally, in order to receive a score of five, which corresponds to stage five of the stage model, respondents had to indicate that their organization had a backup procedure for all data, and that it includes all of the elements listed above.

Procedures Total Score. While six separate questions were used to analyze the Procedures domain, only four scores were produced as some questions or response options were combined to match stage model elements. Once that was done, the domain total score protocol outlined in the introduction to this section was used to calculate the ranges for each domain, and to place respondents into a stage in the procedures domain. The score ranges for the procedures domain are summarized in Table 9 below.

Table 9

Procedure Domain Score Range

	Score Range
Stage 1	4-7
Stage 2	8-11
Stage 3	12-15
Stage 4	16-19
Stage 5	20

Budget

As noted in previous research, budgeted data management activities may lead to more mature data management practices within an organization (Mattia, 2011, p. 130). In addition, researchers have found that a lack of funding for data management practices results in significant challenges, including a lack of access to technology for data tasks and a lack of qualified data management expertise, which could lead to a less developed, or less mature data management program (Beatch et al., 2012; Carman & Fredricks, 2008; Stoecker, 2007). As a result, budgeted data management tasks were important elements to include in the Nonprofit Data Management Stage Model. Two survey questions were used to measure budgeted data management maturity in respondent organizations. The respondents were first asked which data management tasks or items were included in their organizational budget. Respondents were able to select any of the following options:

- New software such as Microsoft Excel or other programs or annual software licenses.
- Hardware such as a computer or server.
- Consulting services for data collection or reporting.
- Data quality audits.
- IT software or hardware expansion.
- Permanent organization-wide budget item for data management tasks.
- Data training for staff that use data regularly.
- Data training for any staff interested in data.
- Program evaluation.

Budget – Question 1 Scoring. In order to give each respondent a score that corresponds to a data management stage, the number of options the respondent selected were summed. However, data training tasks were given more weight, as they are represented as specific elements in the stage model. So, in order to be in a stage higher than one, respondents had to select "data training for staff that use data regularly" and for stage four and higher, respondents had to select "data training for any staff interested in data." This was done because data training is the foundational requirement for staff to be able to manage and use data effectively, and budgeting for this element represents more mature data management budgeting.

For respondents to receive a score of one, which corresponds to stage one of the stage model, they had to select between zero and two response options. For a score of two, which corresponds to stage two, respondents were required to select "data training for any staff that use data regularly” as well as three or fewer responses. To receive a score of three, which corresponds to stage three, respondents were required to select “data training for staff that use data regularly” as well as five to six categories total. In order to receive a score of four, which corresponds to stage four, respondents were required to select “data training for any staff interested in data” and five or more other categories. Respondents who selected “Permanent organization-wide budget item for data management tasks” were automatically scored into stage five with a score of five. This permanent organization-wide budget for data management represents organizations who are most dedicated to data management tasks, and who budget for those tasks annually.

The second question used to measure data management maturity as related to the budget domain asked respondents how their organization, as a whole, viewed spending on data management technology. Respondents could select one of the options below:

- Money should only be spent on technology that can serve multiple purposes.
- Money should be spent on technology only when required for a specific project or task.
- Money should be spent on technology whenever funding is available.
- There should be a budget for required technology.
- There should be a budget for required technology as well as new or innovative technology.

Budget – Question 2 Scoring. Scoring for this question was simpler, as each option were tied directly to elements in the stage model, in ascending order. Respondents who selected “money should only be spent on technology that can serve multiple purposes” were given a score of one, which corresponds to stage one on the stage model. Respondents who selected “money should be spent on technology only when required for a specific project or task" were given a score of two, which corresponds to stage two, and so on.

Budget Total Score. Once the first two combined questions and the third question were scored, the total score was summed. Once that was done, the domain total score protocol outlined in the introduction to this section was used to calculate the ranges for each domain, and to place respondents into a stage in the budget domain. The score ranges for the budget domain are summarized in Table 10 below.

Table 10

Budget Domain Score Range

	Score Range
Stage 1	2-3
Stage 2	4-5
Stage 3	6-7
Stage 4	8-9
Stage 5	10

Stage Model Total Score

Once all domains were scored, the total scores for each domain were summed, and respondents were broken into five groups, using the same procedure for each domain total score. The minimum possible score was 12 and the maximum possible score was 60. By using the same procedure that was used to place respondents into a stage in each domain, the ranges outlined in Table 11 below placed each respondent into a final stage.

Table 11

Stage Model Total Score Range

	Score Range
Stage 1	12-23
Stage 2	24-35
Stage 3	36-47
Stage 4	48-59
Stage 5	60

Scoring – Revised 2-Factor Stage Model

Scoring for the factored data was a simpler process than the data matrix scoring protocol, since the two factors did not correspond directly to the elements in the data matrix. As such, a simplified sum of scores was used to assign a score for each respondent in each of the factors.

Factor One – Culture. The Culture factor included 19 items that were identified using the EFA. The majority of the questions in factor one (12) come from the Culture domain, while the next most common are from the Leadership & Planning domain (4), then Budget (2) and finally there is one question from the Procedures domain.

Of the 19 items, 10 used a simple zero or one score, indicating that a respondent does or does not agree with the indicated statement. Eight of the remaining nine items come from a single matrix-style survey question which asked respondents how their organization spends their time related to data management practices. Respondents could indicate that they never, rarely, sometimes, frequently, or almost always spend time on a list of tasks. Each statement represents increasingly advanced data management practices, and as such proportionally higher scores were given if a respondent indicated that more time was spent on more advanced data management tasks. To assign a proportionally larger score for each statement, the statements were assigned a number from one to eight. Each frequency response was also assigned a number from zero, for never, to four, for almost always. The final score was calculated by multiplying the statement number by the frequency response selected. For example, if a respondent indicated that they sometimes (frequency response assigned a two) spend time making plans on how to use data more effectively (statement assigned a three) would receive a score of 6 for this response (2x3). This scoring mechanism allowed for a proportional increase in scores for data management practices that move from basic to more advanced while also giving weight to the amount of time spent on each of those activities.

The final item was a single-select, multiple choice survey question in the Culture domain which asked which statement most closely matched how an organization is currently using data. Each response statement corresponded to increasingly advanced data management practices and received an increasingly higher score from zero (for the lowest stage) to four (the highest stage).

For the Culture factor, the minimum possible score was zero, and the maximum score possible was 158. These scores were used to test the relationship between data management practice maturity and program outcomes and outputs. Table 12 below displays each survey item, the domain it was originally assigned to, and the possible scores a respondent could receive for that item.

Table 12

Culture Factor Items

Survey Question	Original Domain	Revised Score Range
Does your organization have hardware, such as a computer or server, in its budget?	Budget	0 - 1
Does your organization have IT software or hardware expansion in its budget?	Budget	0 - 1
Does your organization view data as useful for program planning?	Culture	0 - 1
Does your organization view data as useful for budget decisions?	Culture	0 - 1
Does your organization view data as useful for supervision?	Culture	0 - 1
In general, which most closely matches how your organization is currently using data (only as far as it is require by funders/regulations; use data for funders/regulations but recognize that its use can be expanded; use data for funders/regulations but are talking about ways that its use can be expanded; use data for more than just funding requirements and are making plans to expand its use; use data on a regular basis and are currently carrying out plans to expand its use; use data in most aspects of operations and consistently works to make data collection and use more efficient)?	Culture	0 - 4
Does your staff never, rarely, sometimes, frequently, or almost always spend time talking about how to use data to meet reporting requirements?	Culture	0 - 4
Does your staff never, rarely, sometimes, frequently, or almost always spend time talking about the need to use data more effectively?	Culture	0 - 8
Does your staff never, rarely, sometimes, frequently, or almost always spend time making plans on how to use data more effectively?	Culture	0 - 12
Does your staff never, rarely, sometimes, frequently, or almost always spend time carrying out plans on how to use data more effectively?	Culture	0 - 12
Does your staff never, rarely, sometimes, frequently, or almost always spend time in continuous quality improvement about how to manage data better?	Culture	0 - 20
Does your staff never, rarely, sometimes, frequently, or almost always spend time using data to answer questions about or programs and services?	Culture	0 - 24

Does your staff never, rarely, sometimes, frequently, or almost always spend time exploring data to find patterns or information that will help us improve our services?	Culture	0 - 28
Does your staff never, rarely, sometimes, frequently, or almost always spend time researching new or innovative ways to manage data?	Culture	0 - 32
Does your organization have paid training or workshops that are program specific available to staff?	Leadership & Planning	0 - 1
Does your organization have on-the-job training that is not program specific available to staff?	Leadership & Planning	0 - 1
Does your organization have no data training process in place for staff?	Leadership & Planning	0 - 1
There are no data goals in my organization's strategic plan.	Leadership & Planning	0 - 1
Our organization has information about where data should be stored in our data backup procedures.	Procedures	0 - 1

Factor Two – Process. Factor two includes 12 items that were identified using the EFA. Nine of the questions in Process factor come from the Procedures domain, while the next most common are from the Leadership & Planning domain (two), then Budget (one).

Scoring for the Process factor was accomplished similarly to the Culture factor. Four of the questions involved assigning a simple zero or one, with a zero indicating that a respondent did not select the statement, while a one indicated that the respondent did select the statement. The remaining eight questions allowed respondents to indicate that their organization does not have, has for some data, or has for all data, written procedures for eight data management tasks. Scores for each statement were assigned based on if the respondent selected does not have, which was given a score of zero, has for some data, which was given a score of one, and has for all data, which was given a score of three.

For the Process factor, the minimum possible score possible was zero, and the maximum score possible was 20. These scores were used to test the relationship between data management practice maturity and program outcomes and outputs. A summary of all survey items, the original domain, and the revised score range is presented in Table 13 below.

Table 13

Process Factor Items

Survey Question	Original Domain	Revised Score Range
Does your organization have training for any staff interested in data in its budget?	Budget	0 - 1
There goals to improve data use for decision making in my organization's strategic plan.	Leadership & Planning	0 - 1
My organization does not have a strategic plan.	Leadership & Planning	0 - 1
My organization does not have, has for some data, or has for all data, a written procedure for describing the process by which we analyze data.	Procedures	0 - 2
My organization does not have, has for some data, or has for all data, a written procedure for describing guidelines for how we report data.	Procedures	0 - 2
My organization does not have, has for some data, or has for all data, a written procedure for how we define data.	Procedures	0 - 2
My organization does not have, has for some data, or has for all data, a written procedure written procedure for protection against hacking or unauthorized access of data.	Procedures	0 - 2
My organization does not have, has for some data, or has for all data, a written procedure for what to do in case of hacking or unauthorized access of data.	Procedures	0 - 2
My organization does not have, has for some data, or has for all data, a written procedure defining data sharing requirements.	Procedures	0 - 2
My organization does not have, has for some data, or has for all data, a written procedure ensuring data collected is accurate.	Procedures	0 - 2
My organization does not have, has for some data, or has for all data, a written procedure ensuring data is backed up.	Procedures	0 - 2
Our organizations primary motivation behind writing data management policies including wanting to have a document for new staff training.	Procedures	0 - 1

Conclusion

This chapter outlined the research design process, the 50-question survey instrument, and presented the scoring protocol that was used for analysis and testing of the three hypotheses. A description of the research subjects was presented, and a recruitment strategy was outlined. In subsequent chapters, respondent surveys were analyzed and scored and placed into the stage model in each of the four domains, as well as scored using the revised two-factor model. The scores were then used to explore the relationship between the data management maturity of their organization and their reported program outcomes and outputs.

Chapter 4: Research Findings

Introduction

This chapter presents the response rate for the survey, and descriptive statistics of the respondent demographics and organizational demographics, including size, structure, clients served, and budget. Next this chapter presents the frequency of scores for each of the data management domains as well as the stage model total scores and the revised 2-factor model scores. The chapter concludes with a presentation of the results of the regression analyses and Spearman's Rho analysis used to test the three research hypotheses.

Findings

Response Rate

Of the 878 estimated individuals to receive the survey link, 168 individuals started the survey with 73 respondents found to be eligible for the study (they provided domestic violence shelters services to clients) and who answered at least half of the survey questions. The response rate with eligible respondents was calculated to be 8%, lower than is desired for an electronic survey.

Respondent Demographics

A total of 73 respondents completed enough of the survey to be included in the analysis, and 72 provided demographic data about themselves, including the number of years they have worked with the organization they are filling the survey out for, their degree earned (if any), and a list of the formal data training they have received.

Of the 72 respondents who provided demographic data, eight reported not having earned any college degree. The most common degree earned among the remaining respondents was a master's degree or equivalent, with 35 (48.6%) selecting that response. The next most common degree was a bachelors or equivalent, with 27 (37.5%) selecting that response, with the remaining 2 (2.7%) reporting having earned a Ph.D. or equivalent.

Respondents also reported any formal and/or paid training they have received in specialized categories that utilize data, including data analysis, data management, data visualization, research, program evaluation, and strategic planning. The most common data training reported were program evaluation with 26 (36.1%) and strategic planning with 22 (30.6%) of respondents selecting those options. While these two categories were the most commonly selected, it is important to note that all of the skills listed were selected by less than half of the respondents. Table 14 below shows a summary of the results for this survey question.

Table 14

Respondent Training in Data Management

	N	Total (%) Selected	Total (%) Not Selected
Data Analysis	72	12 (16.7%)	60 (83.3%)
Data Management	72	15 (20.8%)	57 (79.2%)
Data Visualization	72	5 (6.9%)	67 (93.1%)
Research	72	17 (23.6%)	55 (76.4%)
Program Evaluation	72	26 (36.1%)	46 (63.9%)
Strategic Planning	72	22 (30.6%)	50 (69.4%)

Respondents were also asked to provide the number of years that they have been employed with the organization that they were reporting for. Responses to this question ranged from less than one year to a maximum of 33 years, with an average of 9.3 years at the organization.

Finally, respondents were asked to name their title in their organization. The majority of respondents (48 or 66.7%) reported being the executive director or CEO. 21 (29.2%) respondents described their role as an assistant director or manager-level position and the remaining three (4.2%) respondents described their position as a coordinator or equivalent.

Organizational Demographics

Budget. Of the 73 respondents, only 69 reported an annual budget for their organization with budget dollar amounts reported as low as $78,030 and has high as $13,000,000. The mean annual budget for the organizations was $1,788,872.

Four additional budget questions were asked to gain a greater understanding of the sources of funding for the DV shelter respondents. Each respondent was asked to report the percentage of their funding that comes from each the following three sources: government grants, private grants, and donations. In addition, respondents were asked to identify what percentage of their budget was flexible, meaning it could be spent on what the organization felt was needed, as opposed to restricted by external requirements.

Most of the funding for organizations came from government grants. An average 69% of funding came from government grants according to respondents, and the minimum amount reported was 35%. Private grants were the next most common, with an average of 15% of funding coming from this source and a minimum of 2%, while donations made up the lowest percentage of funding, with an average of 13% of budgets coming from this source and a minimum of 1%.

Budget flexibility varied more so than budget funding sources, with anywhere between 0% and 85% of budgets being reported as flexible. The average percentage of flexible budgets was reported at 22%. See Table 15 for the minimum, maximum, and average percentage of the budget funding sources and flexibility.

Table 15

Budget by Funding Source

Funding Source	N	Min	Max	Mean
Government Grants	68	35.0%	100.0%	69.5%
Private Grants	61	2.0%	54.0%	15.3%
Donations	63	1.0%	55.0%	13.2%
Budget Flexibility	65	0.0%	85.0%	21.9%

Organization Size. The size and structure of the organization was measured using two survey questions; how many paid employees an organization has, and what type of leadership structure the organization operates under.

The number of paid employees an organization reported varied from between one and five paid employees to over 30 paid employees, with the majority reporting over 30. A breakdown of the number of organizations within each paid employee category can be seen in Table 16 below.

Table 16

Number of Paid Employees

	Frequency	Percent
1-5	3	4.1%
6-10	15	20.5%
11-15	10	13.7%
16-20	8	11.0%
21-30	14	19.2%
31 and over	23	31.5%
Total N	73	

Respondents could select one of four options to describe their leadership structure; a top-down structure with an executive director, department heads, and employees; a non-traditional top-down structure, a non-traditional structure with co-equal leadership responsibilities; or other. One respondent did not answer this question, meaning 72 respondents did reply. None of the respondents reported that their agency utilized a non-traditional leadership structure with co-equal responsibilities. Most respondents, 76% (55), described their administrative structure as utilizing a traditional top-down model, while nearly 21% (15) selected a non-traditional top-down structure. The remaining 2.8% (2) selected "other" as their organizational structure. Table 17 shows the leadership structure frequencies.

Table 17

Leadership Structure

Leadership Structure	Frequency	Percent
Traditional top-down structure with an executive director, department heads, and employees	55	76.3%
We have a non-traditional top-down structure	15	20.8%
Non-traditional structure with co-equal leadership responsibilities	0	0.0%
Other	2	2.8%
Total N	72	

Client Population. The populations that organizations served was also analyzed, with organizations self-selecting their typical clients into three categories: rural clients, urban clients, and a mix of the two. All 73 shelters responded to this question. Most respondents (48%, 35) reported that their clients are rural, while only 11% (8) reported serving urban clients. The remaining 40% (29) reported serving a mix of urban and rural clients. See Table 18 for a breakdown.

Table 18

Client Population by Location

	Frequency	Percent
Rural residents	35	48.6%
Urban residents	8	11.1%
A mix of urban and rural residents	29	40.3%
Total N	72	

Organizational Demographics Association with Outputs and Outcomes

Organizational demographic factors discussed in the previous section, including budget, number of employees (which is a measure for the size of the organization), structure, and location were all included in the logic model that guided this study. Two of those factors proved to be unusable for this study.

The organizational structure survey question showed very little variability in the responses. Because of the lack of variability in the response to this question it was dropped from the organizational demographic analysis.

The second organizational demographic question that was dropped from the analysis was the location of the organization. While the majority of the respondents were split between two options, serving rural and a mix of urban and rural clients, the decision to drop this question was made based on the nature of domestic violence (DV) shelter operations. Since DV shelters are supported by federal funding via state DV coalitions, many urban shelters operate satellite offices in rural areas of the state in which they are funded. As such, the results for this question are potentially artificially skewed toward those that serve rural clients, with only eight respondents reporting working with primarily urban clients. Because of the likelihood of this inaccurate representation of the primary location of the organization, this variable was dropped from the analysis.

The remaining two organizational demographic variables used to explore the association between those demographics and outputs and outcomes were budget and size of the organization.

Budget. The budget survey question asked respondents to report their budget in dollars, and was used to explore the association between budget and the outcome and output variables. None of the analyses proved to be significant but one, the output variable total clients served that are not residing in the shelter. The analysis of variance showed that the effect of budget on total clients served who were not residing at the shelter was significant, $F\ (56, 16) = 2.472$, $p = .024$.

The second demographic used was size of the organization. An analysis of variance showed that the effect of size of an organization on all output variables is significant with the exception of the total number of legal advocacy services provided. Those results are shown in Table 19 below.

Table 19

Organizational Outputs Analysis of Variance

	F	P
Total number of clients served in shelter	(5, 67) 8.696	.000
Total number of clients served not residing in shelter	(5, 67) 2.370	.049
Total legal advocacy services provided	(5, 67) 2.302	.054
Total community presentations provided	(5, 67) 2.663	.030

An analysis of variance showed that the effect of size on the average length of stay in the shelter was not significant, $F\ (5, 67)\ .376$, $p = .863$. An analysis of variance showed that the effect of size on the percent of clients that exit the shelter to stable housing was significant, F (5, 67), 2.481, p = .040.

Scored Questions

As described in the scoring protocol in chapter three, scoring for each domain in the nonprofit data management stage model involved multiple survey questions that were tied to the data management elements (or activities) within that domain. The following section presents the scores organizations received in each of the four domains.

Culture. The culture domain used three survey questions to calculate the score a respondent would receive in this category.

The first was a question asking respondents to identify how, in general, their organization viewed data. Table 20 show a breakdown of the responses for each option below.

Table 20

Culture Survey Question 1 Frequencies

	N	**Total (%) Selected**	**Total (%) Not Selected**
Data is viewed as useful for getting funding	73	71 (97.3%)	2 (2.7%)
Data is viewed as useful for service delivery	73	38 (52.1%)	35 (47.9%)
Data is viewed as useful for program planning	73	65 (89.0%)	8 (11.0%)
Data is viewed as useful for budget decisions	73	60 (82.2%)	13 (17.8%
Data is viewed as useful for supervision	73	50 (68.5%)	23 (31.5%)

Respondents were given a score based on the number of responses they selected (more responses equal a higher score). Table 21 shows the score breakdown for this question.

Table 21

Culture Survey Question 1 Score Frequencies

	Frequency	Percent
1	6	8.2%
2	3	4.1%
3	11	15.1%
4	26	35.6%
5	27	37.0%

The next question used to calculate the culture score asked respondents how their organization is generally using data. The results of that question are in Table 22 below.

Table 22

Culture Survey Question 2 Frequencies

	N	Frequency	Percent
Only as far as it is required by funders or regulations	73	5	6.8%
Use data for funders or regulations but recognize that its use can be expanded	73	14	19.2%
Use data for funders or regulations but are talking about ways that its use can be expanded	73	12	16.4%
Use data for more than funding & making and carrying out plans to expand use	73	20	27.4%
Use data in most aspects of operations and consistently works to make data collection and use more efficient	73	22	30.1%

Each response in this question corresponds directly to a stage in the stage model. A score of 1 to 5 was assigned in ascending order. See Table 23 below for a summary of how many respondents scored into each stage.

Table 23

Culture Survey Question 2 Score Frequencies

	Frequency	Percent
1	5	6.8%
2	14	19.2%
3	12	16.4%
4	20	27.4%
5	22	30.1%

The final scored question for the culture domain asked respondents how often their staff spent time on data management and use tasks. The tasks were presented in a matrix format and were able to select never, rarely, sometimes, frequently, or almost always for each task. A summary of the responses is in Table 24 below.

Table 24

Culture Survey Question 3 Frequencies

		Never	Rarely	Sometimes	Frequently	Almost Always
Our staff spends time talking about how to use data to meet reporting requirements.	N	1	7	27	28	9
	%	1.4%	9.7%	37.5%	38.9%	12.5%
Our staff spends time talking about the need to use data more effectively.	N	2	11	21	30	8
	%	2.8%	15.3%	29.2%	41.7%	11.1%
Our staff spends time making plans on how to use data more effectively.	N	3	14	27	24	4
	%	4.2%	19.4%	37.5%	33.3%	5.6%
Our staff spends time carrying out plans on how to use data more effectively.	N	3	17	30	19	2
	%	4.2%	23.9%	42.3%	26.8%	2.8%
Our staff spends time in continuous quality improvement about how to manage data better.	N	2	15	29	21	4
	%	2.8%	21.1%	40.8%	29.6%	5.6%
Our staff spends time researching new or innovative ways to manage data.	N	7	27	23	13	1
	%	9.9%	30.8%	32.4%	18.3%	1.4%
Our staff spends time using data to answer questions about our programs and services.	N	3	3	22	31	13
	%	4.2%	4.2%	30.6%	43.1%	18.1%
Our staff spends time exploring data to find patterns or information that will help us improve our services.	N	2	9	38	16	6
	%	2.8%	12.7%	53.5%	22.5%	8.5%

Scoring for this question was accomplished by summing the total responses in each category for each respondent. Respondents were given a score between one and five based on the frequency category which has the most responses. If there was a tie, respondents were given the higher of the two scores. A summary of the scores can be found in Table 25 below.

Table 25

Culture Survey Question 3 Score Frequencies

	Frequency	Percent
1	4	5.5%
2	13	17.8%
3	28	38.4%
4	24	32.9%
5	4	5.5%

Culture Total Score. To calculate the total score for the culture domain, the scores for all survey questions in this domain were summed. The minimum score a respondent could receive was three (as there are three questions), and the maximum score is 15. The average total score for the culture domain was 10.6. The total scores are presented in Table 26 below.

Table 26

Culture Domain Total Score Frequencies

	Frequency	Percent
3	2	2.7%
5	1	1.4%
6	2	2.7%
7	2	2.7%
8	8	11.0%
9	9	12.3%
10	11	15.1%
11	8	11.0%
12	9	12.3%
13	12	16.4%
14	6	8.2%
15	3	4.1%

Culture Rank. Once a total score was calculated, each respondent had to be sorted into stages for the culture domain. This was accomplished by placing the respondent into the domain if their score was equal to or less than the score below the minimum score for the next highest stage. A more detailed description of this scoring protocol was outlined in the previous chapter. A summary of the number of respondents sorted into each of the five stages is presented in Table 27 below.

Table 27

Culture Rank Frequencies

	Frequency	Percent
1	3	4.1%
2	12	16.4%
3	28	38.4%
4	27	37.0%
5	3	4.1%

Leadership and Planning. The leadership and planning domain used three survey questions to calculate the score respondents would get in this category.

The first question asked respondents to select the most appropriate response to how their organization works to include data management tasks in staff job descriptions. The majority of respondents selected responses that indicated data management tasks are already in job descriptions for most staff (45.8%) or for all staff (33.3%). A full summary of the responses can be found in Table 28 below.

Table 28

Leadership and Planning Survey Question 1 Frequencies

	N	Frequency	Percent
Data management tasks are not in job descriptions and we are not reviewing or updating	72	5	6.9%
We have talked about reviewing job descriptions to include data management, but we haven't made plans to do so yet	72	5	6.9%
We have made plans or are implementing plans to review data management tasks in job descriptions	72	5	6.9%
Our jobs descriptions already include data management tasks for most employees who manage data	72	33	45.8%
Data management tasks are already in job descriptions for all employees who manage data	72	24	33.3%

Each response in this question corresponds directly to a stage in the stage model. A score of one to five was assigned in ascending order. See Table 29 below for a summary of how many respondents scored into each stage.

Table 29

Leadership and Planning Survey Question 1 Score Frequencies

	Frequency	Percent
1	5	6.9%
2	5	6.9%
3	5	6.9%
4	33	45.8%
5	24	33.3%

The next question in the leadership and planning domain asked respondents to identify the type of data management training that is available to staff. Respondents were able to select any choice that applied to their organization. A summary of the responses is presented in Table 30 below.

Table 30

Leadership and Planning Survey Question 2 Frequencies

	N	Total (%) Selected	Total (%) Not Selected
Paid training or workshops that are program specific	73	21 (28.8%)	52 (71.2%)
Paid training or workshops that are not program specific	73	8 (11.0%)	65 (89.0%)
On-the-job training that is program specific	73	64 (87.7%)	9 (12.3)%
On-the-job training that is not program specific	73	19 (26.0%)	54 (70.0)%
No data training process in place	73	4 (5.5%)	69 (94.5%)

Scoring for this question related directly to elements in the stage model, and respondents were given a score from 1 to 5 based on their selection. Table 31 shows the score frequencies for this question.

Table 31

Leadership and Planning Survey Question 2 Score Frequencies

	Frequency	Percent
1	4	5.6%
2	38	52.8%
3	7	9.7%
4	15	20.8%
5	8	11.1%

The next question in the leadership and planning domain asked respondents about the data management goals that are present in their organizations strategic plan. Respondents were able to select all options that applied to their organization. Table 32 shows the frequencies for each response that was selected.

Table 32

Leadership and Planning Survey Question 3 Frequencies

	N	Total (%) Selected	Total (%) Not Selected
Goals to start data collection	72	9 (12.5%)	63 (87.5%)
Goals to start using data for decision making	72	9 (12.5%)	63 (87.5%)
Goals to improve data collection	72	28 (38.9%)	44 (61.1%)
Goals to improve data use for decision making	72	19 (26.0%)	53 (73.6%)
Goals to expand the use of data beyond current practice	72	26 (36.1%)	46 (63.9%)
Goals to explore new and innovative ways to use data	72	21 (28.8%)	51 (69.9%)
There are no data goals in my organization's strategic plan	72	18 (25.0%)	54 (75.0%)
My organization does not have a strategic plan	72	9 (12.5%)	63 (87.5%)

Scoring for this question related directly to elements in the stage model, and respondents were given a score from one to five based on their selection. Table 33 shows the score frequencies for this question.

Table 33

Leadership and Planning Survey Question 3 Score Frequencies

	Frequency	Percent
3	2	2.7%
5	2	2.7%
6	1	1.4%
7	2	2.7%
8	11	15.1%
9	11	15.1%
10	8	11.0%
11	11	15.1%
12	5	6.8%
13	6	8.2%
14	5	6.8%
15	6	8.2%

Leadership and Planning Total Score. To calculate the total score for the leadership and planning domain, the scores for all survey questions in this domain were summed. The average total score for this domain was 9.6. For the leadership and planning domain, the total scores are presented in Table 34 below.

Table 34

Leadership and Planning Domain Total Score Frequencies

	Frequency	Percent
1	27	37.5%
2	0	0.0%
3	12	16.7%
4	12	16.7%
5	21	29.2%

Leadership and Planning Rank. Once a total score was calculated, each respondent had to be sorted into stages for the leadership and planning domain. This was accomplished by placing the respondent into the domain if their score was equal to or less than the score below the minimum score for the next highest stage. A more detailed description of this scoring protocol was outlined in the previous chapter. A summary of the number of respondents sorted into each of the five stages for the leadership and planning domain is presented in Table 35 below.

Table 35

Leadership and Planning Total Rank Frequencies

	Frequency	Percent
1	5	6.8%
2	24	32.9%
3	24	32.9%
4	17	23.3%
5	3	4.1%

Procedures. The procedures domain used a total of six survey questions for scoring, four scored questions, and two questions that included skip-logic that would determine if the subsequent question were shown to the respondent.

The first question in the procedures domain asked respondents if they had written procedures for a variety of data management tasks. Respondents could indicate if they had procedures for all data, for some data, or that they do not have a procedure for the given data task. A summary of the responses can be found in Table 36 below.

Table 36

Procedures Survey Question 1 Frequencies

		Does Not have	For Some Data	For All Data
Protection against hacking or unauthorized access	N	5	15	39
	%	8.5%	25.4%	66.1%
What to do in the case of hacking or unauthorized access	N	16	13	29
	%	27.6%	22.4%	50.0%
Ensuring data collected is accurate	N	5	21	32
	%	8.6%	36.2%	55.2%
Ensuring data is backed up	N	4	9	46
	%	5.5%	12.3%	63.0%
Describing the process by which we analyze data	N	20	29	10
	%	33.9%	49.2%	16.9%
Describing guidelines for how we report data	N	14	30	14
	%	24.1%	51.7%	24.1%
Defining data sharing requirements	N	7	20	32
	%	11.9%	33.9%	54.2%
How to define data	N	9	30	19
	%	15.5%	51.7%	32.8%
Other Procedures related to data	N	9	24	11
	%	20.5%	54.5%	25.0%

Scoring for this question was based on the number of responses in each of the categories, with more weight given to the option “for all data” and “for some data” than for “does not have.” Respondents were given a score between one and five, which corresponds to the five stages in the stage model. A summary of the frequency and percentages for score can be found in Table 37 below.

Table 37

Procedures Survey Question 1 Score Frequencies

	Frequency	Percent
1	15	20.5%
2	11	15.1%
3	18	24.7%
4	18	24.7%
5	11	15.1%

The next question used to score respondents in the leadership and planning domain asked respondents about their organizations motives in creating data management polices and procedures. Respondents were first given a question that asked if they had polies or procedures that address data management, data sharing, or data use. If they selected “no” or “don’t know” they were automatically given a score of 1 and were not shown the subsequent question. If they responded yes, they were shown the subsequent question. 10 respondents (13.9%) answered “no” and 3 respondents (4.2%) answered “don’t know” and were not shown the subsequent question. For those respondents who were shown the subsequent question, their responses are summarized in Table 38 below.

Table 38

Procedures Survey Question 2 Frequencies

	N	Total (%) Selected	Total (%) Not Selected
Because regulations or laws require it	58	43 (58.9%)	15 (25.9%)
Our organization values data management and use	58	39 (67.2%)	19 (32.8%)
Our organization wanted to have a document for new staff training	58	13 (22.4%)	45 (77.6%)
Our organization wanted to ensure data we collect is done so the same year after year	58	26 (44.8%)	32 (55.2%)

Scoring for this question was directly linked to elements in the stage model. Respondents could receive a score between 1 and 5. A summary of the number of respondents receiving each score can be seen in Table 39 below.

Table 39

Procedures Survey Question 2 Score Frequencies

	Frequency	Percent
1	16	21.9%
2	7	9.6%
3	31	42.5%
4	10	13.7%
5	9	12.3%

The next question in the procedures domain asked respondents how their organization ensured data that they collected and reported is accurate. Respondents could select one of the following: data is not checked for accuracy, data is checked for accuracy when a problem is found, data is checked for accuracy on an occasional basis, and data is checked for accuracy on a regularly scheduled basis. A summary of the responses can be found in Table 40 below.

Table 40

Procedures Survey Question 3 Frequencies

	Frequency	Percent
Data is not checked for accuracy	0	0.0%
Data is checked for accuracy when a problem is found	9	12.3%
Data is checked for accuracy on an occasional basis	15	20.5%
Data is checked for accuracy on a regularly scheduled basis	49	67.1%
Total N	73	

Scoring for this question was done by utilizing responses from the first question in this section. If respondents indicated that they had a policy or procedure for ensuring data collected was accurate, and they checked for accuracy occasionally or regularly, they were given higher scores than those who did not. A summary of the scores can be seen in Table 41 below.

Table 41

Procedures Survey Question 3 Score Frequencies

	Frequency	Percent
1	17	23.3%
2	4	5.5%
3	4	5.5%
4	11	15.1%
5	37	50.7%

The final two survey questions used to score the procedures domain involve the respondent organization's data backup procedures. The first question, only shown to respondents who indicated that they had a backup procedure, asks what data is included in that procedure. 17 (30.9%) of respondents indicated that some of their organizations data was included in their procedures, while 38 (69.1%) indicated that all of their organizations data was included in their procedures. The second question asked respondents what activities were covered in the backup procedure. A summary of those responses can be found in Table 42 below.

Table 42

Procedures Survey Question 4 Frequencies

	N	Total (%) Selected	Total (%) Not Selected
How often data should be backed up	53	47 (88.7%)	6 (11.3%)
Where data should be stored	53	47 (88.7%)	6 (11.3%)
How to perform data backups	53	26 (49.1%)	27 (50.9)

As with the previous survey question, scoring this question was done by combing the responses from this question and the previous one which asked what data was included in the procedure. Respondents were given a higher score for more of the above elements selected combined with having those policies and procedures for most or all data. A summary of the scores can be seen in Table 43 below.

Table 43

Procedures Survey Question 4 Score Frequencies

	Frequency	Percent
1	4	6.8%
2	6	10.2%
3	11	18.6%
4	21	35.6%
5	17	28.8%

Procedures Total Score. To calculate the total score for the procedures domain, the scores for all survey questions in this domain were summed. The average score for this domain was 13.2. For the procedures domain, the total scores are presented in Table 44 below.

Table 44

Procedures Total Score Frequencies

	Frequency	Percent
3	4	5.5%
5	1	1.4%
6	10	13.7%
7	1	1.4%
8	1	1.4%
9	1	1.4%
10	7	9.6%
11	6	8.2%
12	5	6.8%
13	8	11.0%
14	12	16.4%
15	11	15.1%
16	3	4.1%
17	1	1.4%
18	2	2.7%
19	4	5.5%
20	1	1.4%

Procedures Rank. Once a total score was calculated, each respondent had to be sorted into stages for the procedures domain. This was accomplished by placing the respondent into the domain if their score was equal to or less than the score below the minimum score for the next highest stage. A more detailed description of this scoring protocol was outlined in the previous chapter. A summary of the number of respondents sorted into each of the five stages for the procedures domain is presented in Table 45 below

Table 45

Procedures Domain Rank Frequencies

	Frequency	Percent
1	15	20.5%
2	3	4.1%
3	26	35.6%
4	27	37.0%
5	2	2.7%

Budget. The final domain in the stage model involves the organization's budgeted data management practices. A total of three survey questions related to spending on data management practices and equipment were asked to score respondent organizations.

The first question asked respondents to identify which items the organization currently has in its budget. Respondents could select all that applied to their organization. A summary of the responses can be found in Table 46 below.

Table 46

Budget Survey Question 1 Frequencies

	N	Total (%) Selected	Total (%) Not Selected
New software such as Microsoft Excel or other programs or annual software licenses	73	49 (67.1%)	24 (32.9%)
Hardware such as a computer or server	73	49 (67.1%)	24 (32.9%)
Consulting services for data collection or reporting	73	23 (31.5%)	50 (68.5%)
Data quality audits	73	12 (16.4%)	61 (83.6%)
IT software or hardware expansion	73	28 (38.4%)	45 (61.6%)
Permanent organization-wide budget item for data management tasks	73	17 (23.3%)	56 (76.7%)
Data training for staff that use data regularly	73	26 (35.6%)	47 (64.4%)
Data training for any staff interested in data	73	8 (11.0%)	65 (89.0%)
Program evaluation	73	24 (32.9%)	49 (67.1%)

Scoring for this question was accomplished by summing the number of responses selected in addition to weighing some elements higher than others based on their placement in the stage model. A summary of the scores can be seen in Table 47 below.

Table 47

Budget Survey Question 1 Score Frequencies

	Frequency	Percent
1	20	27.4%
2	14	19.2%
3	21	28.8%
4	1	1.4%
5	17	23.3%

The second question that was used to score the budget domain asked respondents how their organization, as a whole, views spending on data management technology. Respondents could select only one of the options. A summary of those responses can be seen in Table 48 below.

Table 48

Budget Survey Question 2 Frequencies

	N	Frequency	Percent
Money should only be spent on technology that can serve multiple purposes	72	21	29.2%
Money should be spent on technology only when required for a specific project or task	72	3	4.2%
Money should be spent on technology whenever funding is available	72	13	18.1%
There should be a budget for required technology	72	17	23.6%
There should be a budget for required technology as well as new or innovative technology	72	18	25.0%

Scoring for this question was directly linked to elements in the stage model. Respondents could receive a score between one and five. A summary of the number of respondents receiving each score can be seen in Table 49 below.

Table 49

Budget Survey Question 2 Score Frequencies

	Frequency	Percent
1	21	29.2%
2	3	4.2%
3	13	18.1%
4	17	23.6%
5	18	25.0%

Budget Total Score. To calculate the total score for the budget domain, the scores for all survey questions in this domain were summed. The average score for this domain was 5.9. For the budget domain, the total scores are presented in Table 50 below.

Table 50

Budget Total Score Frequencies

	Frequency	Percent
2	5	6.9%
3	6	8.3%
4	12	16.7%
5	8	11.1%
6	16	22.2%
7	8	11.1%
8	5	6.9%
9	5	6.9%
10	7	9.7%

Budget Rank. Once a total score was calculated, each respondent had to be sorted into stages for the budget domain. This was accomplished by placing the respondent into the domain if their score was equal to or less than the score below the minimum score for the next highest stage. A more detailed description of this scoring protocol was outlined in the previous chapter. A summary of the number of respondents sorted into each of the five stages for the budget domain is presented in Table 51 below.

Table 51

Budget Domain Rank Frequencies

	Frequency	Percent
1	11	15.3%
2	20	27.8%
3	24	33.3%
4	10	13.9%
5	7	9.7%

Stage Model Total Score. Using the same scoring procedure used for the domain total scores, the stage model total score was calculated by summing all of the domain total scores. The average total score was 39.4. The frequency for all total scores is presented in Table 52 below.

Table 52

Stage Model Total Score Frequencies

	Frequency	Percent
20	1	1.4%
22	1	1.4%
23	1	1.4%
24	1	1.4%
25	3	4.1%
26	1	1.4%
27	1	1.4%
28	1	1.4%
30	2	2.7%
31	4	5.5%
32	2	2.7%
33	1	1.4%
35	2	2.7%
36	4	5.5%
37	2	2.7%
38	7	9.6%
39	6	8.2%
40	3	4.1%
41	2	2.7%
43	2	2.7%
44	3	4.1%
45	2	2.7%
46	3	4.1%
47	2	2.7%
48	3	4.1%
49	2	2.7%
50	3	4.1%
51	1	1.4%
52	1	1.4%
53	4	5.5%
55	1	1.4%
56	1	1.4%

Stage Model Total Rank. By using the minimum score of the total stage model rank through the score just below the minimum score for the next highest domain, each respondent was sorted into a stage for all domains. This is the same procedure used in the domains, however, since stage 5, the highest stage, can only be reached if a respondent scored into stage 5 in all domains, this difficult to reach level was not accomplished by any of the respondents. As such, Table 53 below shows the frequencies for only stages 1-4.

Table 53

Stage Model Total Rank Frequencies

	Frequency	Percent
1	3	4.1%
2	18	24.7%
3	35	47.9%
4	17	23.3%
5	0	0.0%

Instrument Validation

Prior to analyzing the data to explore relationships between data management practices and program outputs and outcomes, an exploratory factor analysis was performed to test the survey instrument validity. While it would be expected that the survey items related to each data management domains would load on one of the four domain factors, in practice, two primary factors emerged with strong loading from multiple items. Some items double-loaded in both factors, and others were eliminated for not reaching a pre-set coefficient absolute value of .320. In addition, further elimination of factor items was based on including only those items with an Eigenvalue greater than one. Once the factor analysis was completed, the survey items went from 55 survey items (10 in the Budget domain, 14 in the culture domain, 14 in the Leadership and Planning domain, and 17 in the Procedures domain) to 31 elements, (19 in factor one, and 12 in factor 2). The majority of items that loaded on the first factor were culture, and the culture items were loaded with values over .500, with 8 over .600. Since the majority of the items were related to the culture domain in the nonprofit data management stage model, factor one was named the Culture factor for clarity.

The second factor from the EFA had questions that primarily related to the procedures domain (9 of the 11 items) and the remaining two items were related to strategic planning. Because this domain contained items related to various processes that occur at an organization, the second factor was named the Process factor for clarity. The rotated component matrix is presented in Table 54 below.

Table 54

Rotated Component Matrix – Revised 2-Factor Stage Model

	Original Domain	Factor 1	Factor 2
Our staff spends time carrying out plans on how to use data more effectively.	Culture	0.814	--
Our staff spends time making plans on how to use data more effectively.	Culture	0.800	--
Our staff spends time exploring data to find patterns or information that will help us improve our services.	Culture	0.790	--
Our staff spends time using data to answer questions about our programs and services.	Culture	0.779	--
Our staff spends time talking about how to use data to meet reporting requirements.	Culture	0.774	--
Our staff spends time in continuous quality improvement about how to manage data better.	Culture	0.771	--
Our staff spends time talking about the need to use data more effectively.	Culture	0.660	--
Our staff spends time researching new or innovative ways to manage data.	Culture	0.629	--
Data is viewed as useful for budget decisions	Culture	0.576	--
In general, which most closely matches how your organization is currently using data?	Culture	0.526	--
Data is viewed as useful for supervision	Culture	0.504	--
Present in Strategic Plan: There are no data goals in my organization's strategic plan	Leadership & Planning	-0.481	--

	Original Domain	Factor 1	Factor 2
Data is viewed as useful for program planning	Culture	0.431	--
Where data should be stored	Procedures	0.366	--
On-the-job training that is not program specific	Leadership & Planning	0.362	--
No data training process in place	Leadership & Planning	-0.360	--
Hardware such as a computer or server	Budget	0.359	--
Paid training or workshops that are program specific	Leadership & Planning	0.355	--
IT software or hardware expansion	Budget	0.347	--
Describing guidelines for how we report data	Procedures	--	0.824
Describing the process by which we analyze data	Procedures	--	0.744
How to define data	Procedures	--	0.735
What to do in the case of hacking or unauthorized access	Procedures	--	0.665
Defining data sharing requirements	Procedures	--	0.648
Protection against hacking or unauthorized access	Procedures	--	0.610
Ensuring data collected is accurate	Procedures	--	0.577
Ensuring data is backed up	Procedures	--	0.469
Data training for any staff interested in data	Budget	--	0.445
Our organization wanted to have a document for new staff training	Procedures	--	0.382
Present in Strategic Plan: Goals to improve data use for decision making	Leadership & Planning	--	0.342
Present in Strategic Plan: My organization does not have a strategic plan	Leadership & Planning	--	-0.341

Once the two factors were identified, a reliability analysis was carried out on the new 2-factor stage model score. Cronbach's alpha showed the new model to reach acceptable reliability, a = .901.

Regression Analysis

A low response rate for the survey was a consistent challenge, and as such a multiple linear regression analysis was performed using both the culture and process factors that made up the revised 2-factor stage model derived from the EFA as well as the original stage model total score which includes all survey items. Both models also include two organizational demographics: organization size (measured by the total number of employees) and the organization's budget.

Hypothesis 1 – Program Outputs

Three outcome questions were included in the survey based on those identified by the Family Violence Prevention and Service Administration. Respondents were asked to provide a whole number in response to the three survey questions. Those outputs include the number of individuals served residential, number of individuals served non-residential, total number of legal advocacy services provided to individuals, and total presentations provided. Table 55 below shows the responses to those output survey questions. As is evidenced by this table, the number of clients served, and the number of services and presentations provided varied widely. For example, the number of clients receiving legal advocacy services had a minimum of 0 to a maximum of 10,000 in a year.

Table 55

Organization Output Frequency

	N	Min	Max	Mean
Clients served who have resided at shelter	73	10	921	232.9
Clients served not residing at shelter	73	22	20,456	1,681.9
Clients receiving legal advocacy services	73	0	10,000	642.4
Number of community presentations provided	73	2	2,093	151.2

Clients served residing in shelter. A multiple linear regression was calculated to predict the number of clients served that resided in the shelter based on the stage model total score, the size of the organization (total number employees) and the budget of the organization. A significant equation was found ($F(3,69) = 13.168$, $p<.001$), with an R^2 of .364. Organizations' predicted number of clients served who resided in shelter is equal to 7.173 – (-.001) (stage model total score) + 2.699 (size) + -.843 (budget), where the stage model total score is in whole numbers, budget is measured in dollars, and size is coded as 1 = 1-5, 2 = 6-10, 3 = 11-15, 4 = 16-20, 5 = 21-30, and 6 = 31+. The total number of clients served who resided at the shelter decreased by .001 clients for each score increase in the stage model total score, increased by 2.699 clients for each increase in size category, and decreased by .843 clients for each increase in budget dollar amount. Only the size of the organization was a significant predictor of the number of clients served that resided in the shelter.

A multiple linear regression was calculated to predict the number of clients served that resided in the shelter based on the culture factor, the process factor, the size of the organization (total number employees) and the budget of the organization. A significant equation was found ($F(4,68) = 9.787$, $p<.001$), with an R^2 of .365. Organizations' predicted number of clients served who resided in shelter is equal to 6.524 – .008 (culture factor) + (-.032) (process factor) + 2.681 (size) + (-.797) (budget), where the culture factor and process factor, are scores in whole numbers, budget is measured in dollars, and size is coded as 1 = 1-5, 2 = 6-10, 3 = 11-15, 4 = 16-20, 5 = 21-30, and 6 = 31+. The total number of clients served who resided at the shelter increased by .008 clients for each score increase in the culture factor, decreased by -.032 for each score increase the process factor, increased by 2.681 clients for each increase in size category, and decreased by -.797 clients for each increase in budget dollar amount. Only the size of the organization was a significant predictor of the number of clients served that resided in the shelter.

Clients served residing out of shelter. A multiple linear regression was calculated to predict the number of clients served that resided out of the shelter based on the stage model total score, the size of the organization (total number employees) and the budget of the organization. A significant equation was found ($F(3,69) = 4.696$, $p<.05$), with an R^2 of .170. Organizations' predicted number of clients served who resided out of the shelter is equal to -.031 – .006 (stage model total score) + .057 (size) + .386 (budget), where the stage model total score is in whole numbers, budget is measured in dollars, and size is coded as 1 = 1-5, 2 = 6-10, 3 = 11-15, 4 = 16-20, 5 = 21-30, and 6 = 31+. The total number of clients served who resided at the shelter increased by .006 clients for each score increase in the stage model total score, increased by .057 clients for each increase in size category, and increased by .386 clients for each increase in budget dollar amount. None of the independent factors, on their own, were significant predictors of the number of clients served that resided out of the shelter.

A multiple linear regression was calculated to predict the number of clients served that resided out of the shelter based on the culture factor, the process factor, the size of the organization (total number employees) and the budget of the organization. A significant equation was found ($F(4,68) = 3.481$, $p<.05$), with an R^2 of .170. Organizations' predicted number of clients served who resided in shelter is equal to -.003 – .002 (culture factor) + .002 (process factor) + .057 (size) + .396 (budget), where the culture factor and process factor, are scores in whole numbers, budget is measured in dollars, and size is coded as 1 = 1-5, 2 = 6-10, 3 = 11-15, 4 = 16-20, 5 = 21-30, and 6 = 31+. The total number of clients served who resided at the shelter increased by .002 clients for each score increase in the culture factor, increased by .002 for each score increase the process factor, increased by .057 clients for each increase in size category, and increased by .396 clients for each increase in budget dollar amount. None of the independent factors, on their own, were significant predictors of the number of clients served that resided out of the shelter.

Total legal advocacy services provided. A multiple linear regression was calculated to predict the legal advocacy services provided based on the stage model total score, the size of the organization (total number employees) and the budget of the organization. A significant equation was found ($F(3,69) = 3.377$, $p<.05$), with an R^2 of .128. Organizations' predicted number of clients served who resided out of the shelter is equal to -2.564 – .009 (stage model total score) + .001 (size) + .711 (budget), where the stage model total score is in whole numbers, budget is measured in dollars, and size is coded as 1 = 1-5, 2 = 6-10, 3 = 11-15, 4 = 16-20, 5 = 21-30, and 6 = 31+. The legal advocacy services provided increased by .009 clients for each score increase in the stage model total score, increased by .001 clients for each increase in size category, and increased by .711 clients for each increase in budget dollar amount. None of the independent factors, on their own, were significant predictors of the number of clients served that resided out of the shelter.

A multiple linear regression was calculated to predict the number of legal advocacy services provided on the culture factor, the process factor, the size of the organization (total number employees) and the budget of the organization. No significant equation was found ($F(4,68) = 2.441$, $p>.05$), with an R^2 of .126. Organizations' predicted number of legal advocacy services provided is equal to -2.417 – .001 (culture factor) + .008 (process factor) + .004 (size) + .717 (budget), where the culture factor and process factor, are scores in whole numbers, budget is measured in dollars, and size is coded as 1 = 1-5, 2 = 6-10, 3 = 11-15, 4 = 16-20, 5 = 21-30, and 6 = 31+. The legal advocacy services provided increased by .001 clients for each score increase in the culture factor, decreased by .008 for each score increase the process factor, increased by

.004 clients for each increase in size category, and decreased by .717 clients for each increase in budget dollar amount. None of the independent factors, on their own, were significant predictors of the number of clients served that resided out of the shelter.

Total community presentations provided. A multiple linear regression was calculated to predict the number of community presentations provided based on the stage model total score, the size of the organization (total number employees) and the budget of the organization. A significant equation was found ($F(3,69) = 2.981$, $p<.05$), with an R^2 of .115. Organizations' predicted community presentations provided is equal to 2.020 – (-.001) (stage model total score) + .165 (size) + (-.159) (budget), where the stage model total score is in whole numbers, budget is measured in dollars, and size is coded as 1 = 1-5, 2 = 6-10, 3 = 11-15, 4 = 16-20, 5 = 21-30, and 6 = 31+. The community presentations provided decreased by .001 clients for each score increase in the stage model total score, increased by .165 clients for each increase in size category, and decreased by .159 clients for each increase in budget dollar amount. Only the size of the organization was significant predictors of the number of community presentations provided by an organization.

A multiple linear regression was calculated to predict the number of community presentations provided on the culture factor, the process factor, the size of the organization (total number employees) and the budget of the organization. A significant equation was found ($F(4,68) = 4.450$, $p<.05$), with an R^2 of .207. Organizations' predicted number of community presentations provided is equal to 1.639 – .005 (culture factor) + (-.034) (process factor) + .168 (size) + (-.127) (budget), where the culture factor and process factor, are scores in whole numbers, budget is measured in dollars, and size is coded as 1 = 1-5, 2 = 6-10, 3 = 11-15, 4 = 16-20, 5 = 21-30, and 6 = 31+. The community presentations provided increased by .005 clients for each score increase in the culture factor, decreased by .034 for each score increase the process factor, increased by .168 clients for each increase in size category, and decreased by .127 clients for each increase in budget dollar amount. The culture factor, process factor, and the size of the organization were all significant predictors of the number of community presentations provided.

Table 56 and 57 presents the multiple linear regression equation, model significance (P), as well as the significance values of each independent variable within the model.

Table 56

Hypothesis 1 – Original Stage Model Total Score Regression Analysis

	Equation	R^2	P Value			
			Model	Total Score	Size	Budget
Clients residing in shelter	F(4,68) = 13.168	.364	.000	.988	.000	.760
Clients residing out of shelter	F(4,68) = 4.696	.170	.005	.479	.444	.188
# Legal advocacy services	F(4,68) = 3.377	.128	.023	.468	.992	.098
# Community presentations	F(4,68) = 2.981	.115	.037	.916	.037	.605

Table 57

Hypothesis 1 – Revised 2-Factor Stage Model Regression Analysis

	LG Equation	R^2	P				
			Model	Culture Factor	Process Factor	Size	Budget
Clients residing in shelter	F(4,68) = 9.787	.365	.000	.733	.809	.000	.774
Clients residing out of shelter	F(4,68) = 3.481	.170	.012	.519	.904	.450	.181
# Legal advocacy services	F(4,68) = 2.441	.126	.055	.776	.697	.973	.098
# Community presentations	F(4,68) = 4.450	.257	.003	.036	.019	.027	.665

Hypothesis 2 – Data-Driven Decision Making

The second hypothesis in this dissertation states that nonprofits at a later stage of nonprofit data management will use data for decision making more than those at an earlier stage of nonprofit data management. To test this hypothesis the survey asked respondents if they felt that their organizations decision making was influenced by the data they collect and analyze. The majority (55 or 76%) said that decisions were somewhat influenced by data they collect and analyze, while an additional 13 (or 18%) or respondents indicated that decisions are influenced primarily by data that they collect and analyze. The remaining 4 respondents indicated that decisions were not influenced by or were only slightly influenced by data collected or analyzed. See Table 58 below for a summary of the responses to this question.

Table 58

Data for Decision Making Frequency

	N	Frequency	Percent
Not influenced by data that we collect and analyze	72	1	1.4%
Influence slightly by data that we collect and analyze		3	4.2%
Influenced somewhat by data that we collect and analyze		55	76.4%
Influenced primarily by data that we collect and analyze		13	18.1%

Results of the Spearman's rho correlation coefficient indicate no significant association between either of the two factors and the use of data for decision making, Culture factor, rs = .169, p = 0.157, Process factor, rs = .176, p = .1139. However, there was a significant, though small to moderate correlation between the stage model total score and the use of data for decision making. rs = .257, p = 0.029.

Hypothesis 3 – Program Outcomes

We asked respondents two outcome questions related to domestic violence services. The first, "what is the average length of stay in your shelter" required respondents to determine the average number of days a shelter resident stayed at the emergency shelter. The responses to this question varied from 2 days to a max of 371 days. The second question asked respondents "what percentage of your clients exit to stable housing." The variability for this response was from 25% to 100% of clients exiting the shelter to stable housing. The results are summarized in Table 59 below.

Table 59

Organizational Outcomes

	N	Min	Max	Mean
Average length of stay in shelter	72	2	371	54.7
Percent of clients who exit to stable housing	68	25%	100%	69%

Average length of stay in shelter. A multiple linear regression was calculated to predict the average length of stay in shelter based on the stage model total score, the size of the organization (total number employees) and the budget of the organization. No significant equation was found (F(3,69) = .490, p>.05), with an R^2 of .021. Organizations' predicted average length of stay in shelter is equal to 8.865 – (-.027) (stage model total score) + .334 (size) + (-.405) (budget), where the stage model total score is in whole numbers, budget is measured in dollars, and size is coded as 1 = 1-5, 2 = 6-10, 3 = 11-15, 4 = 16-20, 5 = 21-30, and 6 = 31+. The average length of stay in shelter decreased by -.027 clients for each score increase in the stage model total score, increased by .334 clients for each increase in size category, and decreased by .405 clients for each increase in budget dollar amount. None of the independent factors, on their own, were significant predictors of the average length of stay in shelter.

A multiple linear regression was calculated to predict the average length of stay in shelter on the culture factor, the process factor, the size of the organization (total number employees) and the budget of the organization. No significant equation was found (F(4,68) = .413, p>.05), with an R^2 of .024. Organizations' predicted the average length of stay in shelter is equal to 7.660 – .007 (culture factor) + (-.047) (process factor) + .292 (size) + (-.371) (budget), where the culture factor and process factor, are scores in whole numbers, budget is measured in dollars, and size is coded as 1 = 1-5, 2 = 6-10, 3 = 11-15, 4 = 16-20, 5 = 21-30, and 6 = 31+. The average length of stay in shelter increased by .007 clients for each score increase in the culture factor, decreased by .047 for each score increase the process factor, increased by .292 clients for each increase in size category, and decreased by -.371 clients for each increase in budget dollar amount. None of the independent factors, on their own, were significant predictors of the average length of stay in shelter.

Percent of clients that exit to stable housing. A multiple linear regression was calculated to predict the percent of clients that exit to stable housing based on the stage model total score, the size of the organization (total number employees) and the budget of the organization. No significant equation was found ($F(3,69) = 2.001$, $p>.05$), with an R^2 of .080. Organizations' predicted the percent of clients that exit to stable housing is equal to -.046 – (-.001) (stage model total score) + .005 (size) + .028 (budget), where the stage model total score is in whole numbers, budget is measured in dollars, and size is coded as 1 = 1-5, 2 = 6-10, 3 = 11-15, 4 = 16-20, 5 = 21-30, and 6 = 31+. The percent of clients that exit to stable housing decreased by .001 clients for each score increase in the stage model total score, increased by .005 clients for each increase in size category, and increased by .028 clients for each increase in budget dollar amount. None of the independent factors, on their own, were significant predictors of the percent of clients that exit to stable housing.

A multiple linear regression was calculated to predict the percent of clients that exit to stable housing on the culture factor, the process factor, the size of the organization (total number employees) and the budget of the organization. A significant equation was found ($F(4,68) = 3.572$, $p<.05$), with an R^2 of .174. Organizations' predicted percent of clients that exit to stable housing is equal to -.094 – .000 (culture factor) + (-.004) (process factor) + .006 (size) + .030 (budget), where the culture factor and process factor, are scores in whole numbers, budget is measured in dollars, and size is coded as 1 = 1-5, 2 = 6-10, 3 = 11-15, 4 = 16-20, 5 = 21-30, and 6 = 31+. The percent of clients that exit to stable housing did not increase for each score increase in the culture factor, decreased by .004 for each score increase the process factor, increased by .006 for each increase in size category, and increased by .030 for each increase in budget dollar amount. Only the process factor was a significant predictor of the percent of clients that exit the shelter to stable housing.

Table 60 and Table 61 present the multiple linear regression equation, model significance (P), as well as the significance values of each independent variable within the regression model used to test hypothesis three. An explanation of these results follows the summary tables.

Table 60

Hypothesis 3 Original Stage Model Total Score Regression Analysis

			P			
	Equation	R^2	Model	Total Score	Size	Budget
Average length of stay in shelter	F(4,68) = .490	.021	.690	.513	.344	.770
Percent exit to stable housing	F(4,68) = 2.001	.080	.122	.356	.480	.362

Table 61

Hypothesis 3 Revised 2-Factor Stage Model Regression Analysis

			P				
	Equation	R^2	Model	Culture Factor	Process Factor	Size	Budget
Average length of stay in shelter	F(4,68) = .413	.021	.790	.587	.479	.412	.790
Percent exit to stable housing	F(4,68) = 3.572	.080	.011	.100	.006	.445	.299

Conclusion

This chapter presented an analysis of the data collected in the 50-question electronic survey that was sent to domestic violence shelters in the United States. As discussed, the response rate was lower than desired at only 8%. The 73 respondents that were found to have completed the survey and who answered most of the survey questions were used in final analysis. Results from the survey, including organizational demographics, reported outputs and outcomes, as well as scores for each question that was used in the four data management themes were presented. The subsequent final chapter will present the summary of the research findings and the discussion of those findings.

Chapter 5: Summary of Findings, Discussion, and Suggestions for Future Research

Introduction

This chapter summarizes the findings related to instrument reliability, the nonprofit data management stage model scoring, the three research hypotheses, as well as the relationship between nonprofit data management and organizational demographics. The limitations are discussed, most notably sample size. The chapter concludes with a discussion of the findings as well as suggestions for future research.

Summary of Findings

Instrument

Testing the survey instrument was particularly challenging due to low sample size. According to Tabachnick and Fidell (2007) there are two primary reasons that an ample sample size is required for a factor analysis. First, correlation coefficients are less reliable with fewer cases, and second, a larger number of factors with lower loadings require more cases for adequate results (p. 613). Since this study only includes 73 viable cases, the reliability of the factor analysis is a significant limitation. However, the factor analysis did produce encouraging results regarding the instrument. The rotated component matrix showed that many of the variables in the culture domain showed moderate to strong correlations, with 12 of the 14 culture variables loading into one factor. In addition, eight of the 12 components showed a loading marker greater than .60. These results suggest that the intent of the instrument, to ask questions that relate to one another and provide an overall picture of data management practices, is on the right track, and with further testing with more data could prove to be a viable tool for nonprofits as they work to improve their data management practices.

Due to the fact that the total number of cases make the reliability estimate of the factor analysis less than poor, it was decided to include both the two factors that emerged from the factor analysis as well as the stage model total score as independent variables in the final analysis. By including results from both the factored components and the stage model total score, future research with larger case numbers can be used to compare results to this initial analysis.

Stage Model Rank

The stage model presented in the literature review involved five distinct stages. The pre-conceptual stage (stage one) includes those organizations that are doing only the minimum required data reporting and are not using data for decision-making. In the conceptualization stage (stage two), organizations are starting to recognize the need to use data for more than just the required reporting and are talking about the need to advance skills and activities. During the preparation stage (stage three), organizations are planning for, and starting to carry out plans to increase data use activities, and they are starting to include data tasks in their strategic plan. In the dedicated management stage (stage four) organizations have started to use data regularly for decision-making, and job description and policy documents reflect organized data goals and tasks and data training for staff is becoming more commonplace. And finally, organizations within the strategic management stage (stage five) are using data to inform more operations within the organization. Strategic goals reflect not only data use but data improvement and exploration of innovative ways to use data. Data procedures are complete, and a robust security system is in place to protect data.

The challenges associated with this stage model, as the literature review revealed, is that nonprofits are not very advanced in terms of data management practices, are not adequately trained to manage and use data, and that funding for such activities in the nonprofit world are lacking (Carman & Fredericks, 2010; Carman & Fredericks, 2008; Stoecker, 2007). So, while the hope would be that the sample population for this study would show a range of maturity levels for data management, the reality proved untrue, if expected. Very few organizations (3) fell within the lowest pre-conceptual stage, and none managed to score highly enough in all domains to qualify for the strategic management stage. As such, using the raw total score was more useful for this analysis. Future research will benefit from a larger sample size from a variety of nonprofits in order to include a wider range of data management practices.

Hypothesis 1

The first hypothesis in this research states that nonprofits at a later stage of the nonprofit data management stage model will have increased outputs compared with those at an earlier stage. Testing this hypothesis was done by using transformed program outputs, including the total number of clients served that resided at the shelter, the total number of clients served that did not reside in the shelter, the total number of legal advocacy services provided, and the total number of community presentations provided as dependent variables.

When using the two factors as the independent variable, as well as the transformed total budget and organization size measured in number of paid employees, the number of clients residing in shelter and the total number of community presentations showed a statistically significant relationship with some independent variables for both the revised 2-factor model as well as the original stage model total score. For both models, the number of clients residing in shelter was only significantly related to the organizations size. For the revised 2-factor stage model both factors and the organizations size were significant predictors of the number of community presentations provided, while for the original stage model total score, only the organizations size significantly predicted the number of community presentations.

The organizations size was a significant predictor of the number of clients residing in shelter and the number of community presentations provided for both models, which is not a surprising result. One would expect the number of outputs to increase as the number of paid employees increases. The culture factor and process factors also significantly predicting the number of community presentations provided, which is encouraging for the revised 2-factor stage model, but additional data will be required to confirm this significant relationship.

Given the low sample size and the lack of significant results between the stage model total score and the dependent variables, and between the revised 2-factors for all but one of the dependent variables, the null hypothesis could not be rejected for hypothesis 1.

Hypothesis 2

To test hypothesis 2, that nonprofits at a later stage of the nonprofit data management stage model will use data for decision making more than those at an earlier stage, we asked survey respondents the following question: "Would you say that decision making in your organization is: not influenced by data that we collect and analyze, influence slightly by data that we collect and analyze, influenced somewhat by data that we collect and analyze, and influenced primarily by data that we collect and analyze."

There is a statistically significant positive relationship between an organization's use of data for decision making and their stage model total score at the .05 level. However, the correlation is considered small to moderate with a correlation coefficient of .247. Additional confirmation with more cases may be needed before the null hypothesis could be rejected. However, this weak correlation may be due to a number of factors.

First, the use of data for decision making could be impacted by other factors, such as the pressure to meet external reporting requirements for funders. As researchers have suggested, most nonprofit organizations use data primarily to satisfy external reporting requirements (Botcheva, White, & Huffman, 2012) and that seems to be true for this sample as well (p. 428). When asked what most closely matches how the respondent's organization is currently using data, less than 31% report regularly using data beyond what is required by external reporting. 31 respondents (42.4%) reported using data only to meet external reporting requirements. 14 of those 31 (19.2%) reported recognizing a need to expand the use of data, while 12 (16.4%) reported actually talking about expanding data use within the organization. The remaining 5 of the 31 reported not talking about expanding data use. 20 respondents (27.4%) reported being in the planning phase or beginning to carry out plans to use data beyond external reporting requirements. That leaves only 22 (30.1%) of organizations using data regularly for more than what is required by external funders or regulations.

It is possible that this emphasis on meeting the demands of external funders or regulations takes precedence over desires of an organization to use data for decision-making, as data collection and reporting are time-consuming activities. In addition, the data collected for external reporting is not always viewed as the most useful data for organizations to make internal decisions (Mitchell & Berlan, 2016; Carnochan, Samples, Myers, & Austin, 2013; Carman & Fredericks, 2010' Stoecker, 2007).

Second, it is possible that a stronger relationship would be present if there were more respondents with more varied stage model total scores. As stated earlier, none of the respondents managed to gain the high score which would have put them in stage 5 of the nonprofit data management stage model, and this lack of advanced data management practices could have impacted this result significantly.

Finally, it is possible that the survey question itself was not sufficient to explore the topic of data use for decision-making. Only one survey question was used, and decision-making can be a very complex process. Exploring this hypothesis could have been more robust with multiple survey questions exploring decision-making from a variety of lenses (funding decisions, programmatic decisions, personnel management decisions, etc.). Given that a statistically significant, albeit weak, relationship exists with the small sample size, this hypothesis deserves further attention in future research.

Hypothesis 3

Exploring the third hypothesis, that nonprofits at a later stage of the nonprofit data management stage model will have better program outcomes than those at an earlier stage, was the most challenging, specifically because of the DV population used for this study. The primary function of DV nonprofits is to provide a safe space for DV survivors who are in acute crisis, meaning their primary service is, from an evaluation perspective, an output (clients served), as opposed to an outcome. While many organizations work toward returning survivors to a place of safety and health, this is not always the focus of the DV shelter itself, but instead other arms of the organization or partners in the community (such as local healthcare facilities, housing authorities, educational/vocational training organizations, etc.). However, our study was limited to only DV shelters. Due to these challenges, only two identified outcomes were used for this study, a client's length of stay in shelter, and the percentage of clients that exit the shelter to stable housing.

Of the two program outcomes, only one, the percent of clients that exit to stable housing, showed a statistically significant relationship, and that relationship was with the process factor that were created after performing the EFA. However, this result should not be overlooked.

While, again, sample size is a serious limitation to this study, the potential relationship that nonprofit data management practices may have on the number of clients who exit acute crisis situations into stability could be very impactful to this population. Again, rejecting the null hypothesis in this study may be premature. While the process factor was found to be related to percent of clients who exited the shelter to stable housing, additional data could strengthen that result and reveal more statistically significant relationships in the original stage model and with the average length of time a client stays at the shelter.

Organizational Demographics

While very few statistically significant associations between the stage model score and the two identified factors were found with this study, the results of the analysis did confirm a common assumption; an organizations size is significantly related to at least two program outputs (number of clients residing in shelter and the number of community presentations provided). This result makes logical sense as more employees should lead to increased services for clients.

This result, while expected, does reinforce the importance of the relationship between staff and program outputs and outcomes. If those employees are better trained in data management, are more amenable to a culture that emphasizes data management and use, are given appropriate policies and procedures related to data management and are given access to data management tools and practices that are better funded; perhaps, they could then provide the most meaningful data driven services to their clients.

The second analyzed demographic, budget, was not significantly related to the any of the dependent variables. Unlike the previously discussed result, this result is somewhat counterintuitive. One would assume that better-funded organizations could tackle more advanced tasks than those that were less funded. However, it is more likely that the funding is restricted in such a way that, even with a larger budget, the percentage of that budget that is dedicated to data management practices remains too small to make a positive difference. According to Carman and Fredericks (2008) only a small percentage of survey clients in their study had data-heavy activities funded through their grants. Since DV shelters rely heavily on grant funds, mostly from government sources, similar restrictions on spending can be expected. A more detailed exploration of organizational budgets in the future could illuminate the nature of the relationship between program budget and data management maturity.

Limitations

As mentioned, many times in this study, the sample size is the most severe limitation encountered. With only 73 useable cases to study, the reliability of the psychometric analysis of the instrument is poor. In addition, the small number of cases could have had a negative impact on the variability in total stage model scoring, leaving no cases in the top scoring stage.

In addition to sample size, the survey instrument itself is a limitation. Compromises had to be made between length of the survey and the variety of questions asked. For example, as described above, only one survey question asked how organization use data for decision-making, where multiple survey questions would have explored this subject more thoroughly. In addition, given the exploratory nature of this study, the limitation of a web-survey was apparent. While the survey questions allowed for a scoring protocol that enabled a quantitative exploration of data management practices, the survey questions lacked the depth and richness that qualitative data could provide. Using a combination of qualitative and quantitative data may allow for a more detailed study of specific data management practices and concepts.

Missing data was also a limitation in this research, especially given the low sample size. Using mean substation to manage missing data allowed for the inclusion of additional respondents, but this method can artificially decrease the variation in responses, and likewise result in lower t-values (Hawthorne & Elliott, 2005, p. 588). However, the percent of responses that required mean substitution were small, typically less than 10%.

Finally, a lack of guiding research can be considered a limitation in this study. Nonprofit data management practices are rarely looked at in research, and when they are, the research is focused on specific data activities, such as program evaluation, as opposed to the management of data itself. This proved to be a significant challenge to overcome, especially when designing a survey instrument.

Discussion and Suggestions for Future Research

Exploring the data management practices of nonprofits is a challenging endeavor. There was very little research on nonprofit data management practices to draw from, and the concept of data management itself, especially in the nonprofit world, is abstract and ever-changing as technology and methodologies evolve. However, this study sought to explore the relationship between nonprofit data management practices and program outputs and outcomes, while also filling a significant gap in literature, and that was accomplished at least to a certain degree. In addition, this research provided opportunity to explore the importance of nonprofit data use and allowed for the following recommendations for nonprofit practitioners and funders, as well as recommendations for future research to explore this topic further.

Recommendations for Nonprofit Data Management Practices

Use Data More. While nonprofits collect a large amount of data, only a small percentage of that data is used, and a minority of organizations in this study actually used the data they collect for decision making. Since funders, especially those from government sources, typically require organizations to provide evidence that their programs are having a positive impact on the clients they serve, it is increasingly important that nonprofits use the data they collect to plan and improve their programming, and to manage their personnel, their budgets, and their activities.

In addition, data can be used to expand programming and to effectively advocate for increased funding for programming that might not be already in place. Data that is available, such as secondary data sources (community surveys, census data), or organizational data (social media data, building use statistics) can be leveraged in grant applications that, if funded, could be used to start new programs based on new community needs. This data can also be presented to current funders to justify budget modifications or funding increases to meet increased demand or changing populations demographics. However, the need for new programming or increased funding cannot be discovered if the information that exists is not properly analyzed and leveraged. This lack of data use should be viewed as a missed opportunity for nonprofit organizations, and managers should work to increase the use of data in order to increase, improve, or diversify offerings for their clients and their communities.

Manage Data Better. As stated above, data is being collected by nonprofits, but data that is not managed well, that is not discoverable by analysts or software, or that is too disorganized to analyze in a reasonable amount of time or with minimal analytical skill, is not helpful for decision making. Properly managed data should be both secure and accessible and should exist in a format that is useable. However, managing data in this way takes skills and time that nonprofit organizations tend to lack. In this study, none of the 73 organizations scored high enough in the data management domains to be placed into the strategic management stage, meaning none are managing and using data to its fullest strategic potential. A goal of nonprofit organizations should be to manage data as an invaluable resource for generating funds, planning programming, innovating practice, and improving the lives of the clients they serve.

Fund Data Management. While data collection and management practices spoken of above point toward further action that can be taken by nonprofit managers, it is important to consider the funding required to take such action, and the reality that these funds are typically lacking. Data management, and even data itself, is incredibly complex and requires specific training and skills that are in short supply in nonprofit organizations. While funders are increasingly requiring data-informed practice and reporting, they are not typically providing overhead funding that would finance those activities or the personnel who can perform them. This starvation cycle, as described by Gregory & Howard (2009), occurs when funders unrealistic expectations about the costs of managing an organization, including providing adequate infrastructure, technology, and training, such as data management or analysis training, results in nonprofits not allocating enough funding to these essential tasks. As organization continue to skimp on this essential overhead, funders continue to expect low overhead percentages, thereby continuing the cycle of inadequate funding (p. 49). As this research shows, the size of the budget does not translate into more advanced data management practices or data or use, and this may be due to restrictive nature of the funding itself. If nonprofits are to become more proficient in data management and data use, funders will need to start providing funding for technology, training, analysts, and researchers to realize that goal.

Recommendations for Future Research

Larger Sample Size and More Information. Since the sample size is a significant limitation of this study, future research should be focused on further testing of the instrument and the stage model by collecting more responses from a variety of nonprofit organizations. This study also revealed a consistent lack of quality information about how nonprofits are actually managing, storing, and using their data. Further exploratory research, potentially using qualitative methods such as focus groups or in-depth interviews coupled with analysis of existing organizational documents such as strategic plans and evaluations, could prove significantly useful in refining the stage model.

More Varied Nonprofits. The sample population, while taken from nonprofits from around the USA in 23 different states, is limited to those that provide emergency shelter to domestic violence survivors in acute crisis. This is a very specific nonprofit organization with limited services, which is not typical of all organizations. Nonprofit organizations, while operating under similar tax guidelines, are incredibly varied in their size, structure, services, clients served, and primary mission. If the stage model is to be used to guide nonprofit data management practices, additional data from a variety of nonprofits will be required to ensure the model is both general enough to encompass nonprofits of all types, but specific enough as to be useful for assessment and data management planning. Controlling for additional mission areas will be required for this research. The inclusion of existing models and research related to various mission areas (such as fundraising and advocacy) would allow for exploration of the proposed model's usefulness in assessing the data management maturity of a variety of nonprofits.

Additional Instrument Testing. The nonprofit data management stage model was an attempt to synthesize the research on data management practices in the nonprofit world. The resulting model can, with additional testing, be simplified and improved upon to act as a guide that enables nonprofits to recognize the linear path from limited to advanced data management practices. The survey instrument, while complex, is a first step toward developing a tool that can be used by nonprofit managers to assess their data management stage which can provide a starting point on the road toward strategic data management. Additional testing of the original survey with more participants as well as testing of a revised and simplified instrument should be a goal of future research.

Additional Exploration of Budget. While very few statistically significant associations between the stage model score and the revised 2-factor stage model were found with this study, the results of the analysis did confirm a common assumption: organization size is related to program outputs. Results from this study also challenged the assumption that budget is related to data management practices, outputs and outcomes. This lack of a relationship between budget and data management maturity warrants future study. A more detailed understanding of the relationship between funding and data management could lead to recommendations to program funders in the future. These evidence-based recommendations are essential if support for these important activities is to ever grow.

Conclusion

This study sought to explore the relationship between nonprofit data management practices and program outputs and outcomes. Due to small sample size, additional research will be needed before conclusions about the viability of the nonprofit data management stage model and the survey instrument can be definitively drawn. However, this study has contributed to a significant gap in the literature on the data management practices of nonprofit organizations by providing a synthesized model that maps out specific data tasks and how those tasks may change as an organization's data management practices mature over time.

Results from this study confirm that organizational size (number of employees) is a significant predictor of some program outputs, and that future research could focus on equipping those employees with data management tools to study the impact that this preparation has on said outputs and outcomes.

While some statistically significant relationships between program outputs and outcomes and the revised 2-factor stage model were present, none were strong enough to enable the researchers to confidently reject any of the null hypotheses. However, the presence of statistically significant results in each of the three hypotheses is encouraging and justifies continued exploration of this topic in future research.

Finally, the lack of a statistically significant relationship between an organizations budget and data management maturity may indicate a need for funding that is less restrictive and allows for necessary technology, training, or personnel trained in data management for nonprofits. If data management practices and data use are to increase, funders must consider allocating grant dollars specifically for needed technology, personnel, and skills that are currently lacking in the nonprofit world.

Appendix A

Survey Instrument

EXEMPT RESEARCH STUDY
INFORMATION SHEET
Department of Public Policy and Leadership

TITLE OF STUDY: Nonprofit Data Management Stage Theory
INVESTIGATOR(S) AND CONTACT PHONE NUMBER: Ashley Hernandez-Hall – 775-209-6220 & Dr. Patricia Cook-Craig – 702-972-1059

The purpose of this study is to understand how nonprofit organizations use and manage information. You are being asked to participate in the study because you meet the following criteria: You work for or manage a nonprofit organization that collects, stores, analyzes, and/or reports data.

If you volunteer to participate in this study, you will be asked to do the following: Complete an online, confidential survey. No identifying information about you will be shared.

This study includes only minimal risks. The study will take 30 *minutes* of your time. You *will not* be compensated for your time.

For questions regarding the rights of research subjects, any complaints, or comments regarding the manner in which the study is being conducted you may contact **the UNLV Office of Research Integrity – Human Subjects at 702-895-2794, toll free at 877-895-2794, or via email at IRB@unlv.edu.**

Your participation in this study is voluntary. You may withdraw at any time. If you have questions about this study, please feel free to email Ashley Hall at ashley.hernandez-hall@unlv.edu.

Participant Consent:

By clicking “yes” below and completing this survey, you are acknowledging that you have reviewed this consent form and are agreeing to participate in this evaluation. If you do not wish to volunteer to participate, please check “no” below.

Organizational Demographics

To the best of your ability, please respond to the following questions about your organization.

1. About how many paid employees does your organization have?
    a) 1-5
    b) 6-10
    c) 11-15
    d) 16-20
    e) 20-30
    f) 30+

1. When thinking about your administrative space, or the place that daily administrative operations occur, which best fits your organization?
    a) Our administration operates out of one building
    b) Our administration operates out of several separate buildings
    c) Our administration is mainly remote with a building for meetings or other in-person events

1. When considering your organizational structure, which best fits your organization?
    a) We have a traditional top-down structure with an executive director, department heads, and employees
    b) We have a non-traditional top-down structure with an executive director and employees below with equal levels of responsibility
    c) We have a non-traditional organizational structure with co-equal leadership responsibilities
    d) We have another type of organizational structure, please describe:

1. How would you describe the populations you serve?
    a) Urban residents
    b) Rural residents
    c) Mix of urban and rural

1. What is the annual operating budget for your organization?

1. About what percentage of your operating budget comes from the following:

a) Federal or state grants
b) Private or other grants
c) Donations
d) Fees to clients
e) Other (please describe)

The following four questions will ask about the outputs your organization reports. Please use the most recent complete reporting year for this data.

In the last year….

1. How many individuals have you served that have resided at your shelter?
1. How many individuals have you served that have not resided at your shelter?
1. How many individuals have received legal advocacy services?
1. How many community presentations have you provided?

The next two questions will be related to the program outcomes you report. Please use the most recent complete reporting year for this data.

In the last year….

1. What is the average length of time clients stay in your shelter?

1. What percentage of clients exit to stable housing, by stable housing we mean housing that is a safe transitional or permanent place to live?

Now I am going to ask you some questions about your position in the organization. Please answer the following questions as they relate to your job and position in the organization.

1. What is your position in the organization?

1. How long have you been employed at your organization?

1. Do you have a degree? What is it?

1. Do you have formal training (paid training, certificates, degree) in any of the following?
   a) Data analysis
   b) Data management
   c) Data visualization
   d) Research
   e) Program evaluation
   f) Strategic planning

1. Do you have any other training related to data management, data collection or data reporting? Please describe.

Culture

Now I would like to ask you questions about your perceptions of the ways in which your organization uses information.

1. How often does your organization **collect data** for the following entities?

Entity	Very often	Somewhat often	Not very often	Rarely	Never
External funder (grant funder, government organization)	☐	☐	☐	☐	☐
Board of directors	☐	☐	☐	☐	☐
Executive director	☐	☐	☐	☐	☐
Program directors/supervisors	☐	☐	☐	☐	☐
Outside agencies	☐	☐	☐	☐	☐
Clients	☐	☐	☐	☐	☐
Public	☐	☐	☐	☐	☐

1. How often does your organization **report data** to the following entities?

Entity	Very often	Somewhat often	Not very often	Rarely	Never

External funder (grant funder, government organization)	☐	☐	☐	☐	☐
Board of directors	☐	☐	☐	☐	☐
Executive director	☐	☐	☐	☐	☐
Program directors/supervisors	☐	☐	☐	☐	☐
Outside agencies	☐	☐	☐	☐	☐
Clients	☐	☐	☐	☐	☐
Public	☐	☐	☐	☐	☐

1. How often are reports or evaluations created by your organization shared with the following?

Entity	Very often	Somewhat often	Not very often	Rarely	Never
External funder (grant funder, government organization)	☐	☐	☐	☐	☐
Board of directors	☐	☐	☐	☐	☐
Executive director	☐	☐	☐	☐	☐
Program directors/supervisors	☐	☐	☐	☐	☐
Front-line workers/staff in your organizations	☐	☐	☐	☐	☐
Politicians or political entities	☐	☐	☐	☐	☐
	Very often	Somewhat often	Not very often	Rarely	Never
Trainings or conference attendees	☐	☐	☐	☐	☐
Outside agencies	☐	☐	☐	☐	☐
Clients	☐	☐	☐	☐	☐

Public	☐	☐	☐	☐	☐

1. Would you say that decision making in your organization is:
   a) Influenced primarily by data that we collect and analyze
   b) Influenced somewhat by data that we collect and analyze
   c) Influence slightly by data that we collect and analyze
   d) Not influenced by data that we collect and analyze

1. In general, how **does your organization as a whole** view data (select all that apply):
   a) Data is viewed as useful for getting funding
   b) Data is viewed as useful for service delivery
   c) Data is viewed as useful for program planning
   d) Data is viewed as useful for budget decisions
   e) Data is viewed as useful for supervision

1. In general, which most closely matches how your organization is currently using data?
   a) Only as far as it is required by funders or regulations
   b) Use data for funders or regulations but recognize that its use can be expanded
   c) Use data for funders or regulations but are talking about ways that its use can be expanded
   d) Use data for more than just funding requirement and are making plans to expand its use
   e) Use data on a regular basis and are currently carrying out plans to expand its use
   f) Use data in most aspects of operations and consistently works to make data collection and use more efficient

1. When considering the ways in which data is stored in your organization, which of the following is most accurate:
   a) Data is primarily stored with people who are responsible for collecting it (such as on individual computers or in separate databases).
   b) Data is primarily stored centrally (such as on a shared server or database).
   c) Data is stored both centrally and with the people who are responsible for collecting it

1. How accessible would you say that data is to the following staff members?

Staff	Open Access	Accessible by request	Not accessible

Board of directors	☐	☐	☐
Executive director	☐	☐	☐
Program directors/supervisors	☐	☐	☐
Front-line workers/staff in your organizations	☐	☐	☐

Leadership and Planning

Now I would like to ask you questions about how the leaders in your organization mange data and how they assign others to manage data.

1. How would you characterize the general approach to data management in your organization?
   a) Data is primarily the responsibility of program director/supervisor
   b) Data is primarily the responsibility of a data analyst
   c) Data is primarily the responsibility of the executive director or other administrator
   d) Data is primarily the responsibility of the person in charge Information Technology
   e) There is no one person who is primarily responsible for data in my organization
   f) Other (please specify):

1. How are data management tasks assigned for each of the following positions within your organization?

Staff	Data management tasks are specified in job description	Data management tasks are expected but not specified in job descriptions	Data management tasks are not expected	This position does not exist in my organization
Executive Director or upper management	☐	☐	☐	☐
Information Technology (IT) specialist	☐	☐	☐	☐
Data analyst	☐	☐	☐	☐

Temporary employee or contractor	☐	☐	☐	☐
Project directors/supervisors	☐	☐	☐	☐
Front-line workers/staff within your agency	☐	☐	☐	☐
Volunteers	☐	☐	☐	☐
Other position not listed (please specify)	☐	☐	☐	☐

1. Which of the following best describes how your organization ensures job descriptions includes data management tasks?
   a) Data management tasks are not in job descriptions and we are not reviewing
   b) We have talked about reviewing job descriptions to include data management, but we haven't made plans to do so
   c) We have made plans to review data management tasks in job descriptions and have begun to do so
   d) Our jobs descriptions already include data management tasks for most employees who manage data and there are no plans to update
   e) Data management tasks are already in job descriptions for all employees who manage data

Now we are going to ask you questions about how people in your organization are trained in data management and how your organization plans for data management tasks.

1. Which of the following forms of data training are available to staff in your organization (select all that apply)?
   a) Paid training or workshops that are program specific
   b) Paid training or workshops that are not program specific
   c) On-the-job training that is program specific
   d) On-the-job training that is not program specific
   e) No data training process in place

1. You indicated that your organization provides paid training and workshops on data management, please indicate who in your organization is eligible for that training.

a) Any staff who desire it
b) Only staff who use data regularly as a part of their job description
c) Other (please specify):

1. How would you characterize your organizations data management skill level as a whole? – move up to before 30 – change responses to minimal to where they need to be for the five stages
a) Our data management skills are less than where they need to be
b) Our data management skills are almost where they need to be
c) Our data management skills are where they need to be

1. How would you characterize your organizations approach to improving data management skills?
a) We are not working on improving the data management skills of staff
b) We are training our staff to be able to meet the data management needs of the organization
c) We are offering data management training beyond what is required and encourage continuous education in data management

The following questions are about how your staff spends time talking about or planning data management tasks.

1.

Staff	Never	Rarely	Sometimes	Frequently	Almost always
Our staff spends time talking about how to use data to meet reporting requirements	☐	☐	☐	☐	☐
Our staff spends time talking about the need to use data more effectively	☐	☐	☐	☐	☐
Our staff spends time making plans on how to use data more effectively	☐	☐	☐	☐	☐
Our staff spends time carrying out plans on how to use data more effectively	☐	☐	☐	☐	☐
Our staff spends time in continuous quality improvement about how to manage data better	☐	☐	☐	☐	☐

Our staff spends time researching new or innovative ways to manage data.	☐	☐	☐	☐	☐
Our staff spends time using data to answer questions about our programs and services	☐	☐	☐	☐	☐
Our staff spends time exploring data to find patterns or information that will help us improve our services	☐	☐	☐	☐	☐

1. Which of the following are present in your organizations strategic plan?
   a) Goals to start data collection
   b) Goals to start using data for decision making
   c) Goals to improve data collection
   d) Goals to improve data use for decision making
   e) Goals to expand the use of data beyond current practice
   f) Goals to explore new and innovative ways to use data
   g) There are no data goals in my organizations strategic plan
   h) My organization does not have a strategic plan

Procedures

The following questions ask about the procedures that are in place around data management within your organization.

1. Does your organization have policies and/or procedures that address any of the following: data management, data sharing, or use?

1. What is included in your data sharing policies and procedures? Check all that apply
   a) Who can access data
   b) How to access data
   c) What data can be shared
   d) Who data can be shared with
   e) How data can be shared
   f) Other: Please specify
   g) Our organization does not have a policies or procedures about sharing data

1. Please select whether your organization does not have, has for some data, or has for all data, any of the following written procedures.

Procedure Type	Do not have	For some data	For all data
Protection against hacking or unauthorized access	☐	☐	☐
What to do in the case of hacking or unauthorized access	☐	☐	☐
Ensuring data collected is accurate	☐	☐	☐
Ensuring that data is collected	☐	☐	☐
Ensuring data is backed up	☐	☐	☐
Describing the process by which we analyze data	☐	☐	☐
Describing guidelines for how we report data	☐	☐	☐
Defining data sharing requirements	☐	☐	☐
How to define data (for example, a written description of a piece of information that is collected such as a count of unduplicated individuals to receive a service)	☐	☐	☐
Other Procedures (please specify)	☐	☐	☐

1. What would you say was your organizations motivation behind writing data management policies? – Check all that apply
   a) Because regulations or laws require it.
   b) Our organization values data management and use
   c) Our organization wanted to have a document for new staff training
   d) Our organization wanted to use policies to ensure compliance
   e) Our organization wanted to ensure data we collect is done so the same year after year
   f) Other: Please explain

1. You indicated that your organization has a data backup procedure, what data is currently included in that procedure:

a) All data
b) Some data
c) No data
d) Unsure

1. You indicated that your organization has a data backup procedure, please select each of the following are in it:
   a) How often data should be backed up
   b) Where data should be stored
   c) How to perform data backups
   d) Other: please specify

1. How does your organization, as a whole, ensure data collected and reported is accurate?
   a) Data is not checked for accuracy
   b) Data is checked for accuracy when a problem is found
   c) Data is checked on an occasional basis
   d) Data is checked for accuracy on a regularly scheduled basis

Budget

The next several questions are related to your organizations budget.

1. About what percentage of your budget is flexible, meaning it can be spent on what you feel is needed as opposed to restricted by external requirements?

1. Which of the following does your organization have in its budget?
   a) Software such as a cloud server or hardware such as a computer
   b) Consulting services for data collection or reporting
   c) Data quality audits
   d) IT server or hardware expansion
   e) Permanent organization-wide budget item for data management tasks
   f) Data training for staff that use data regularly
   g) Data training for any staff interested in data
   h) Program evaluation

i) Other budgeted data tasks or items (please specify)

1. How does your organization, as a whole, view spending on data management technology?
   a) Money should only be spent on technology that can serve multiple purposes
   b) Money should be spent on technology only when required for a specific project or task
   c) Money should be spent on technology whenever funding is available
   d) There should be a budget for required technology
   e) There should be a budget for required technology as well as new or innovative technology

Thank you (Qualtrics system generated closing statement)

Appendix B

Complete Reference List

Aiken, P., Allen, M., Parker, B., & Mattia, A. (2007). Measuring data management practice maturity: A community's self-assessment. *Computer*, *40*(4), 42-50. : http://togetherwepass.co.za/http://togetherwepass.co.za/wp-content/uploads/filebase/certified_office_manager/management/study_notes_/Extreme-Information-Management%20Extreme.pdf

Adrian, M. (2011). *Information management goes extreme: The biggest challenges for 21st century CIOs* [Power Point Slides]. SAS: Business Analytics and Intelligence Software.

Altman, M., Wood, A., O'Brien, D., Vadhan, S., & Gasser, U. (2015). Towards a modern approach to privacy-aware government data releases. *Berkeley Technology Law Journal*, *30*(3), 1967-2072.

Andersen, & Henriksen. (2006). E-government maturity models: Extension of the Layne and Lee model. *Government Information Quarterly*, *23*(2), 236-248.

Andrei, K., Pope, E., Hart, A., & Quinn, L. S. (2012). *The state of data in the nonprofit sector*. NTEN. https://act.everyaction.com/2016-nonprofit-data-whitepaper

Arnold, M. E. (2006). Developing evaluation capacity in extension 4-H field faculty: A framework for success. *American Journal of Evaluation*, *27*(2), 257-269.

Arzberger, P., Schroeder, P., Beaulieu, A., Bowker, G., Casey, K., Laaksonen, L., Moorman, D., Uhlir, P. &. Wouters, P. (2004). An international framework to promote access to data. Science. *American Association for the Advancement of Science.*, *303*(5665), 1777-1778.

Baker, C. K., Billhardt, K. A., Warren, J., Rollins, C., & Glass, N. E. (2010). Domestic violence, housing instability, and homelessness: A review of housing policies and program practices for meeting the needs of survivors. *Aggression and Violent Behavior*, *15*(6), 430-439.

Beath, C., Becerra-Fernandez, I., Ross, J., & Short, J. (2012). Finding value in the information explosion. *MIT Sloan Management Review*, *53*(4), 18. https://search.proquest.com/docview/1023761998.

Bedsworth, W., Gregory, A. G., & Howard, D. (2008). Nonprofit overhead costs: Breaking the vicious cycle of misleading reporting, unrealistic expectations, and pressure to conform. *The Bridgespan Group*. https://www.bridgespan.org/insights/library/pay-what-it-takes/nonprofit-overhead-costs-break-the-vicious-cycle

Benjamin, L. (2008). Bearing more risk for results: Performance accountability and nonprofit relational work. *Administration & Society*, *39*(8), 959-983.

Benjamin, L. (2008). Account space: How accountability requirements shape nonprofit practice. *Nonprofit and Voluntary Sector Quarterly*, *37*(2), 201-223.

Bennett, L., Riger, S., Schewe, P., Howard, A., & Wasco, S. (2004). Effectiveness of hotline, advocacy, counseling, and shelter services for victims of domestic violence. *Journal of Interpersonal Violence, 19*(7), 815-829. doi:10.1177/0886260504265687

Berlin, N., Masaoka, J., & Schumann, M. (2017). Two-legged stool: New findings from California on nonprofits and overhead. *Nonprofit Policy Forum, 8*(2), 165-181.

Bernard, C., & Pukstas, K. (2010). *Nonprofit IT staffing: 2009 IT staffing and spending report*. https://www.nten.org/article/2008-it-staffing-spending-report/

Bertot, J., & Choi, H. (2013). Big data and e-government: Issues, policies, and recommendations. *Proceedings of the 14th Annual International Conference on Digital Government Research*, 1-10.

Bloom, N., Lemos, R., Sadun, R., Scur, D., & Van Reenen, J. (2016). International data on measuring management practices. *American Economic Review*, *106*(5), 152-156.

BoardSource (2003). The Sarbanes Oxley Act and implications for nonprofit organizations. *Guidestar*. Washington, DC. https://www.guidestar.org/Articles.aspx?path=/rxa/news/articles/2003/sarbanes-oxley-act-and-implications-for-nonprofit-organizations.aspx

Botcheva, L., White, C. R., & Huffman, L. C. (2002). Learning culture and outcomes measurement practices in community agencies. *American Journal of Evaluation*, *23*(4), 421-434. doi:10.1177/109821400202300404

Boyne, G. (2002). Public and private management: What's the difference? *Journal of Management Studies*, *39*(1), 97-122.

Brohman, Parent, Pearce, & Wade. (2000). *The business intelligence value chain: Data-driven decision support in a data warehouse environment: An exploratory study. System Sciences, 2000*. *Proceedings of the 33rd Annual Hawaii International Conference.* 10 pp. vol.1.

Carman, J. G. (2007). Evaluation practice among community-based organizations: Research into the reality. *American Journal of Evaluation, 28*(1), 60-75.

Carman, J. G. (2009). Nonprofits, funders, and evaluation. *The American Review of Public Administration, 39*(4), 374-390. doi:10.1177/0275074008320190

Carman, J. (2010). The accountability movement: What's wrong with this theory of change? *Nonprofit and Voluntary Sector Quarterly*, *39*(2), 256-274.

Carman, J. (2011). What you don't know can hurt your community: Lessons from a local United Way. *Nonprofit Management and Leadership*, *21*(4), 433-448.

Cameron, K. S., & Quinn, R. E. (2011). *Diagnosing and changing organizational culture: Based on the competing values framework*. John Wiley & Sons.

Carman, J. G., & Fredericks, K. A. (2010). Evaluation capacity and nonprofit organizations. *American Journal of Evaluation, 31*(1), 84-104. doi:10.1177/1098214009352361

Carman, J. G., & Fredericks, K. A. (2008). Nonprofits and evaluation: Empirical evidence from the field. *New Directions for Evaluation, 2008*(119), 51-71. doi:10.1002/ev.268

Carman, J., Leland, S., & Wilson, A. (2010). Crisis in leadership or failure to plan? *Nonprofit Management and Leadership, 21*(1), 93-111.

Carnochan, S., Samples, M., Myers, M., & Austin, M. (2013). Performance measurement challenges in nonprofit human service organizations. *Nonprofit and Voluntary Sector Quarterly, 43*(6), 1014-1032. doi:10.1177/0899764013508009

Cassell, Kay Ann, & Mercado, Marina I. (1999). High performance nonprofit organizations: Managing upstream for greater impact. *The Bottom Line, 12*(1), 45. https://doi.org/10.1108/bl.1999.17012aad.001

Cerchione, R., Esposito, E., & Spadaro, M. R. (2015). A literature review on knowledge management in SMEs. *Knowledge Management Research & Practice, 14*(2), 169-177.

Clarke, R. (2016). Big data, big risks. *Information Systems Journal*, *26*(1), 77-90.

Coe, C., & Brunet, J. (2006). Organizational report cards: Significant impact or much ado about nothing? *Public Administration Review, 66*(1), 90-100.

Cong, X., & Pandya, K. (2003). Issues of knowledge management in the public sector. *Electronic Journal of Knowledge Management, 1*(2), 25-32.

Connolly, P., York, P. (2003). Building the capacity of capacity builders. *The Conservation Company.* https://www.issuelab.org/resources/8470/8470.pdf

Connolly, P., & York, P. (2002). Evaluating capacity-building efforts for nonprofit organizations. *OD Practitioner*, *34*(4), 33-39.

Cupoli, P., Earley, S., & Henderson, D. (2014). DAMA-DMBOK2 Framework. *DAMA International*.

Dahlberg, T., & Nokkala, T. (2015). A framework for the corporate governance of data - theoretical background and empirical evidence. *Business, Management and Education*, *13*(1), 25-45.

De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of big data based on its essential features. *Library Review, 65*(3), 122-135. doi:10.1108/LR-06-2015-0061

Dicke, L. A. (2002). Ensuring accountability in human services contracting: Can stewardship theory fill the bill? *American Review of Public Administration*, *32*(4), 455-470.

Dixon, N., & Shofer, M. (2006). Struggling to invent high-reliability organizations in health care settings: Insights from the field. *Health Services Research*, *41*(42), 1618-1632.

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META Group Research Note*, *6*(70), 1.

Ebrahim, A. (2002). Information struggles: The role of information in the reproduction of NGO-funder relationships. *Nonprofit and Voluntary Sector Quarterly*, *31*(1), 84-114. doi:10.1177/0899764002311004

Eisenhardt, K. M, & Martin, J. A. (2000). Dynamic capabilities: What are they? *Strategic Management Journal*, *21*(10/11), 1105–1121. https://doi.org/10.1002/1097-0266(200010/11)21:10/113.0.CO;2-E

Ekta Arora. (2011). Knowledge management in public sector. *Researchers World*, *2*(1), 165. https://search.proquest.com/docview/1006467877

Farah, B. (2017). A value based big data maturity model. *Journal of Management Policy and Practice, 18*(1), 11-18.

Farkas, M. G., Hinchliffe, L. J., & Houk, A. H. (2015). Bridges and barriers: Factors influencing a culture of assessment in academic libraries. *College & Research Libraries, 76*(2), 150-169. doi:10.5860/crl.76.2.150

Fredericksen, P., & London, R. (2000). Disconnect in the Hollow State: The pivotal role of organizational capacity in community-based development organizations. *Public Administration Review*, *60*(3), 230-239.

Gamage, P. (2016). New development: Leveraging 'big data' analytics in the public sector. *Public Money & Management, 36*(5), 385-390. doi:10.1080/09540962.2016.1194087

Gantz, J., & Reinsel, D. (2011). Extracting value from chaos. *IDC IView*, 1-12.

Garven, S., Hofmann, M., & McSwain, D. (2016). Playing the numbers game. *Nonprofit Management and Leadership*, *26*(4), 401-416.

Gillenson, M. L. (1991). Database administration at the crossroads: The era of end-user-oriented, decentralized data processing. *Journal of Database Management (JDM)*, *2*(4), 1-11.

Gillenson, M. (1982). The state of practice of data administration-1981. *Communications of the ACM*, *25*(10), 699-706.

Gillenson, M. (1985). Trends in data administration. *MIS Quarterly*, *9*(4), 317.

Gonzalez, J. J., & Sawicka, A. (2002, October). A framework for human factors in information security. In *Wseas international conference on information security*, Rio de Janeiro (pp. 448-187).

Grabowski, L., Neher, C., Crim, T., & Mathiassen, L. (2015). Competing values framework application to organizational effectiveness in voluntary organizations: A case study. *Nonprofit and Voluntary Sector Quarterly*, *44*(5), 908-923.

Gray, J. (1996). Data management: Past, present, and future. *IEEE Computer*, *29*(10). http://arxiv.org/abs/cs/0701156

Greenberg, R., & Nunamaker, T. (1987). A generalized multiple criteria model for control and evaluation of nonprofit organizations. *Financial Accountability & Management*, *3*(3-4), 331-342.

Gregory, A. G., & Howard, D. (2009). The nonprofit starvation cycle. *Stanford Social Innovation Review*, *7(*4), 49. https://search.proquest.com/docview/217173506

Gustafson, T., & Fink, D. (2013, August). *Winning within the data value chain.* https://www.innosight.com/insight/winning-within-the-data-value-chain/

Hager, M., Wilson, S., Pollak, T., & Rooney, P. (2003). Response rates for mail surveys of nonprofit organizations: A review and empirical test. *Nonprofit and Voluntary Sector Quarterly*, *32*(2), 252-267.

Hawthorne, G., & Elliot, P. (2005). Imputing cross-sectional missing data: Comparisons of common techniques. *Australian and New Zealand Journal of Psychiatry*, 583-590.

Heimovics, R., & Herman, R. (1993). Executive leadership and resource dependence in nonprofit organizations: A frame analysis. *Public Administration Review*, *53*(5), 419.

Henderson, J., & Venkatraman, N. (1999). Strategic alignment: Leveraging information technology for transforming organizations. *IBM Systems Journal*, *38*(2), 472-484.

Hendricks, M., Plantz, M., Pritchard, K., Carman, Joanne G., & Fredericks, Kimberly A. (2008). Measuring outcomes of United Way–funded programs: Expectations and reality. *New Directions for Evaluation*, *2008*(119), 13-35.

Hikmet, & Chen. (2003). An investigation into low mail survey response rates of information technology users in health care organizations. *International Journal of Medical Informatics*, *72*(1), 29-34.

Hoefer, R. (2000). Accountability in action? Program evaluation in nonprofit human service agencies. *Nonprofit Management and Leadership*, *11*(2), 167-178.

Hoole, E., Patterson, T., Carman, Joanne G., & Fredericks, Kimberly A. (2008). Voices from the field: Evaluation as part of a learning culture. *New Directions for Evaluation*, *2008*(119), 93-113.

Hota, C., Upadhyaya, S., & Al-Karaki, J. (2015). Advances in secure knowledge management in the big data era. *Information Systems Frontiers*, *17*(5), 983-986.

Houchin, S. (2002). Holding ourselves accountable: Managing by outcomes in girls incorporated. *Nonprofit and Voluntary Sector Quarterly*, *31*(2), 271.

Huffman, D., Thomas, K., & Lawrenz, F. (2008). A collaborative immersion approach to evaluation capacity building. *American Journal of Evaluation*, *29*(3), 358-368.

Jacobus, R. (2016). Who will pay for data? *Stanford Social Innovation Review*, *14*(2), 59-60.

Janssen, M., Voort, H., & Veenstra, A. (2015). Failure of large transformation projects from the viewpoint of complex adaptive systems: Management principles for dealing with project dynamics. *Information Systems Frontiers*, *17*(1), 15-29.

Janssen, & Van Den Hoven. (2015). Big and Open Linked Data (BOLD) in government: A challenge to transparency and privacy? *Government Information Quarterly*, *32*(4), 363-368.

Jones, K., & Mucha, R. (2014). Sustainability assessment and reporting for nonprofit organizations: accountability “for the Public Good”. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, *25*(6), 1465-1482.

Jos, P., & Tompkins, M. (2004). The accountability paradox in an age of reinvention: The perennial problem of preserving character and judgment. *Administration & Society*, *36*(3), 255-281.

Khatri, V., & Brown, C. (2010, Jan 1,). Designing data governance. *Communications of the ACM*, *53*, 148-152. doi:10.1145/1629175.1629210 http://dl.acm.org/citation.cfm?id=1629210

Kim, D. Y., & Grant, G. (2010). E-government maturity model using the capability maturity model integration. *Journal of Systems and Information Technology*, *12*(3), 230-244.

Kiron, D. (2013). Organizational alignment is key to big data success. *MIT Sloan Management Review*, *54*(3), N/a.

Klievink, B., Romijn, B., Cunningham, S., & Bruijn, H. (2017). Big data in the public sector: Uncertainties and readiness. *Information Systems Frontiers*, *19*(2), 267-283.

Klievink, & Janssen. (2009). Realizing joined-up government - Dynamic capabilities and stage models for transformation. *Government Information Quarterly*, *26*(2), 275-284.

Lakos, A., & Phipps, S. E. (2004). Creating a culture of assessment: A catalyst for organizational change. *Libraries and the Academy*, *4*(3), 345-361.

Lavine, M., Jules, C., & Good, D. (2014). Paradoxical leadership and the competing values framework. *The Journal of Applied Behavioral Science, 50*(2), 189-205.

Lee, C., & Clerkin, R. (2017). Exploring the use of outcome measures in human service nonprofits: Combining agency, institutional, and organizational capacity perspectives. *Public Performance & Management Review*, *40*(3), 601-624.

Lee, M. R., & Lan, Y. (2011). Toward a unified knowledge management model for SMEs. *Expert Systems with Applications, 38*(1), 729-735. doi:10.1016/j.eswa.2010.07.025

Lenczner, M., & Phillips, S. (2012). From stories to evidence: How mining data can promote innovation in the nonprofit sector. *Technology Innovation Management Review, 2(*7), 10-15. https://search.proquest.com/docview/1614473198

Levitin, A. V., & Redman, T. C. (1998). Data as a resource: Properties, implications, and prescriptions. *Sloan Management Review, 40*(1), 89. https://search.proquest.com/docview/1302987002

Light, P. C., Hubbard, E. T., Kibbe, B., Patrizi, P., Sherwood, K., & Spector, A. (2002). *The capacity building challenge*. Human Interaction Research Institute.

Lucas, J., Henry, & Sutton, J. (1977). The stage hypothesis and the s-curve: Some contradictory evidence. *Communications of the ACM, 20*(4), 254-259. doi:10.1145/359461.359472

Luen, T., & Al-Hawamdeh, S. (2001). Knowledge management in the public sector: Principles and practices in police work. *Journal of Information Science, 27*(5), 311-318.

Lyon, E., Lane, S., & Menard, A. (2008). *Meeting survivors' needs: A multi-state study of domestic violence shelter experiences*. Hartford: School of Social Work, University of Connecticut.

M+R. (2017). *2017 benchmarks study*. https://mrbenchmarks.com/

Malik, P. (2013). Governing big data: Principles and practices. *IBM Journal of Research and Development, 57*(3/4), 1:13. doi:10.1147/JRD.2013.2241359

Mancini, J., Marek, L., Byrne, R., & Huebner, A. (2004). Community-based program research. *Journal of Community Practice*, *12*(1-2), 7-21.

Marco, D. (2003). Meta data management and enterprise architecture. *DM Review, 13*(7), 128.

Margetts, H., & Sutcliffe, D. (2013). Addressing the policy challenges and opportunities of "Big data". *Policy & Internet, 5*(2), 139-146.

Mattia, A. (2011). An exploratory study on using stage theory to develop A data management measurement framework. *International Journal of Management & Information Systems (IJMIS), 15*(4), 123. doi:10.19030/ijmis.v15i4.5810

McAdam, R., & Reid, R. (2000). A comparison of public and private sector perceptions and use of knowledge management. *Journal of European Industrial Training*, *24*(6), 317-329.

Millesen, J., Carman, J., & Bies, A. (2010). Why engage?: Understanding the incentive to build nonprofit capacity. *Nonprofit Management and Leadership*, *21*(1), 5-20.

Mitchell, G. E., & Berlan, D. (2016). Evaluation and evaluative rigor in the nonprofit sector. *Nonprofit Management and Leadership*, *27*(2), 237-250. doi:10.1002/nml.21236

Mott, A. (2006). Hand in hand: Evaluation and organizational development. *Nonprofit Quarterly*, *13*(1), 42-47.

Murray, Vic, & Cutt, James. (2002). *Accountability and effectiveness evaluation in non-profit organizations. In Accountability and effectiveness evaluation in non-profit organizations (Vol. 2).* Routledge.

Murray, V. (2001). The state of evaluation tools and systems for nonprofit organizations. *New Directions for Philanthropic Fundraising*, *2001*(31), 39-50.

Naccarella, Pirkis, Kohn, Morley, Burgess, & Blashki. (2007). Building evaluation capacity: Definitional and practical implications from an Australian case study. *Evaluation and Program Planning*, *30*(3), 231-236.

Ndoye, A. (2010). Creating and sustaining a culture of assessment. *Planning for Higher Education*, *38*(2), 28-39.

Newcomer, K. (2004). How might we strengthen evaluation capacity to manage evaluation contracts? *American Journal of Evaluation*, *25*(2), 209-218.

Nolan, R. L. (1979). Managing the crises in data processing. *Harvard Business Review*, *57*, 126. https://search.proquest.com/docview/227788117

Nolan, R. (1973). Managing the computer resource: A stage hypothesis. *Communications of the ACM*, *16*(7), 399-405. doi:10.1145/362280.362284

NTEN. (2016). *The state of nonprofit data*. https://www.nten.org/knowledge/reports-and-resources/

NTEN. (2012). *The state of nonprofit data*. https://www.nten.org/knowledge/reports-and-resources/

Ottinger, G., & Zurer, R. (2011). Drowning in data. *Issues in Science and Technology*, *27*(3), 71-82.

Parackal, M. (2003). Internet-based and mail survey: A hybrid probabilistic survey approach. *In Proceedings of the 2003 Australian Web Conference*. http://ausweb.scu.edu. au/aw03/papers/parackal.html (accessed April 15, 2004).

Parsons, L., Pryor, C., & Roberts, A. (2017). Pressure to manage ratios and willingness to do so: Evidence from nonprofit managers. *Nonprofit and Voluntary Sector Quarterly*, *46*(4), 705-724.

Peterson, R. (2004). Crafting information technology governance. *Information Systems Management*, *21*(4), 7-22.

Phippen, A., Raza, A., Butel, L., & Southern, R. (2011). Impacting methodological innovation in a local government context: Data sharing rewards and barriers. *Methodological Innovations*, *6*(1), 58-72.

Plantz, M. C., Greenway, M. T., & Hendricks, M. (1997). Outcome measurement: Showing results in the nonprofit sector. *New Directions for Evaluation, 1997*(75), 15-30.

Poole, D., Davis, J., Reisman, J., & Nelson, J. (2001). Improving the quality of outcome evaluation plans. *Nonprofit Management and Leadership*, *11*(4), 405-421.

Prasetyo, H. N., & Surendro, K. (2015). Designing a data governance model based on soft system methodology (SSM) in organization. *Journal of Theoretical and Applied Information Technology*, *78*(1), 46.

Premkumar, G. (2003). A meta-analysis of research on information technology implementation in small business. Jour*nal of Organizational Computing and Electronic Commerce*, *13*(2), 91-121.

Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, *29*(4), 443-459.

Quinn, Robert E, & Cameron, Kim. (1983). Organizational life cycles and shifting criteria of effectiveness: Some preliminary evidence. *Management Science*, *29*(1), 33–51. https://doi.org/10.1287/mnsc.29.1.33

Quinn, Robert E, & Rohrbaugh, John. (1983). A spatial model of effectiveness criteria: Towards a competing values approach to organizational analysis. *Management Science,* *29*(3), 363–377. https://doi.org/10.1287/mnsc.29.3.363

Redman, T. (1998). The impact of poor data quality on the typical enterprise. *Communications of the ACM*, *41*, 79-82. doi:10.1145/269012.269025 http://dl.acm.org/citation.cfm?id=269025

Robey, D., Im, G., & Wareham, J. (2008). Theoretical foundations of empirical research on interorganizational systems: Assessing past contributions and guiding future directions. *Journal of the Association for Information Systems*, *9*(9), 497-518.

Rosenbaum, S. (2010). Data governance and stewardship: Designing data stewardship entities and advancing data access. *Health Services Research*, *45*(52), 1442-1455.

Scholl, H. J., & Scholl, M. C. (2014). Smart governance: A roadmap for research and practice. *iConference 2014 Proceedings*.

Schuh, & Leviton. (2006). A framework to assess the development and capacity of non-profit agencies. *Evaluation and Program Planning*, *29*(2), 171-179.

Sheehan, Kim. (2001). E-mail survey response rates: A review. *Journal of Computer-Mediated Communication*, *6*(2).

Shiri, A. (2013). Linked data meets big data: A knowledge organization systems perspective. *Advances in Classification Research Online*, *24*(1), 16-20.

Sloan, M. (2009). The effects of nonprofit accountability ratings on donor behavior. *Nonprofit and Voluntary Sector Quarterly*, *38*(2), 220-236.

Snibbe, A. C. (2006). Drowning in data. *Stanford Social Innovation Review, 4*(3), 39-45.

Sowa, J., Selden, S., & Sandfort, J. (2004). No longer unmeasurable? A multidimensional integrated model of nonprofit organizational effectiveness. *Nonprofit and Voluntary Sector Quarterly*, *33*(4), 711-728.

Stevenson, Florin, Mills, & Andrade. (2002). Building evaluation capacity in human service organizations: A case study. *Evaluation and Program Planning*, *25*(3), 233-243.

Stoecker, R. (2007). The research practices and needs of non-profit organizations in an urban center. *Journal of Sociology & Social Welfare*, *34*(4), 97.

Stoecker, R. (2006). Neighborhood data systems: A best practice analysis. *Community Development, 37*(4), 109-122.

Sullivan, C. M. (2012). Domestic violence shelter services: A review of the empirical evidence. *Domestic Violence Evidence Project*. http://www.dvevidenceproject.org

Sullivan, C. M. (2011). Evaluating domestic violence support service programs: Waste of time, necessary evil, or opportunity for growth? *Aggression and Violent Behavior, 16*(4), 354-360. doi:10.1016/j.avb.2011.04.008

Susan Phillips, & Michael Lenczner. (2012). From stories to evidence: How mining data can promote innovation in the nonprofit sector. *Technology Innovation Management Review*, 10-15.

Tambe, Prasanna. (2014) Big data investment, skills, and firm value. *Management Science*, *60*(6), 1452-1469.

Tekiner, F., & Keane, J. A. (2013, October). Big data framework. *In 2013 IEEE international conference on systems, man, and cybernetics* (pp. 1494-1499). IEEE.

Thompson, N., Ravindran, R., & Nicosia, S. (2015). Government data does not mean data governance: Lessons learned from a public sector application audit. *Government Information Quarterly, 32*(3), 316-322. doi:10.1016/j.giq.2015.05.001

Tong, Y., & Arvey, R. (2015). Managing complexity via the Competing Values Framework. *Journal of Management Development*, *34*(6), 653-673.

United States. Government Accountability Office, issuing body. (2013). *Federal information security : Mixed progress in implementing program complements; improved metrics needed to measure effectiveness.* Report to Congressional committees.

US Department of Health and Human Services [USDHHS]. (2012). *Navigating the family violence prevention and services program*. https://www.acf.hhs.gov/fysb/resource/fvpsa-state-guide

United Way (2000). *Agency experiences with outcome measurement.* Alexandria, Virginia.

Vilminko-Heikkinen, R., & Pekkola, S. (2017). Master data management and its organizational implementation. *Journal of Enterprise Information Management*, *30*(3), 454-475.

Vinzant, Douglas H., & Vinzant, Janet C. (1996). Strategy and organizational capacity: Finding a fit. Public *Productivity and Management Review*, *20*(2), 139-157.

Walker, G., & Grossman, J. (1999). Philanthropy and outcomes. Dilemmas in a quest for accountability. *Public Private Ventures Briefs (P/PV),* 449-460.

Wang, R., & Strong, D. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, *12*(4), 5-33.

Weber, K., Otto, B., & Österle, H. (2009). One size does not fit all: A contingency approach to data governance. *Journal of Data and Information Quality (JDIQ), 1*(1), 1-27. doi:10.1145/1515693.1515696

Wende, K. (2007). A model for data governance: Organizing accountabilities for data quality management. *ACIS 2007 Proceedings*, 80.

Wells, P. (2012). The non-profit sector and its challenges for governance. *Journal of leadership, accountability, and ethics, 9*(2), 83-93.

Whitaker, G. P., Altman-Sauer, L., & Henderson, M. (2004). Mutual accountability between governments and nonprofits: Moving beyond "surveillance" to "service". *American Review of Public Administration*, *34*(2), 115-133.

Williams, A., & Taylor, P. (2013). Resolving accountability ambiguity in nonprofit organizations. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, *24*(3), 559-580.

Wing, K. T. (2004). Assessing the effectiveness of capacity-building initiatives: Seven issues for the field. *Nonprofit and Voluntary Sector Quarterly, 33*(1), 153-160. doi:10.1177/0899764003261518

Wing, K. (2004). Getting what we pay for: Low overhead limits nonprofit effectiveness. *Urban Institute*. http://webarchive.urban.org/publications/311044.html

Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS quarterly*, 17-41.

Yee-Ching, L. (2004). Performance measurement and adoption of balanced scorecard: A survey of municipal governments in the USA and Canada. *The International Journal of Public Sector Management*, *17*(2/3), 204-221.

Yu, T., & Wu, N. (2009). A review of study on the competing values framework. *International Journal of Business and Management*, *4*(7), 37.

Yun, G., & Trumbo, C. (2000). Comparative response to a survey executed by post, e-mail, & web Form. *Journal of Computer-Mediated Communication*, *6*(1).

References

Andrei, K., Pope, E., Hart, A., & Quinn, L. S. (2012). *The state of data in the nonprofit sector*. NTEN. https://act.everyaction.com/2016-nonprofit-data-whitepaper

Baker, C. K., Billhardt, K. A., Warren, J., Rollins, C., & Glass, N. E. (2010). Domestic violence, housing instability, and homelessness: A review of housing policies and program practices for meeting the needs of survivors. *Aggression and Violent Behavior*, *15*(6), 430-439.

Beath, C., Becerra-Fernandez, I., Ross, J., & Short, J. (2012). Finding value in the information explosion. *MIT Sloan Management Review, 53*(4), 18. https://search.proquest.com/docview/1023761998.

Bernard, C., & Pukstas, K. (2010). *Nonprofit IT staffing: 2009 IT staffing and spending report*. https://www.nten.org/article/2008-it-staffing-spending-report/

Botcheva, L., White, C. R., & Huffman, L. C. (2002). Learning culture and outcomes measurement practices in community agencies. *American Journal of Evaluation, 23*(4), 421-434. doi:10.1177/109821400202300404

Carman, J. G., & Fredericks, K. A. (2010). Evaluation capacity and nonprofit organizations. *American Journal of Evaluation, 31*(1), 84-104. doi:10.1177/1098214009352361

Carman, J. G., & Fredericks, K. A. (2008). Nonprofits and evaluation: Empirical evidence from the field. *New Directions for Evaluation, 2008*(119), 51-71. doi:10.1002/ev.268

Carnochan, S., Samples, M., Myers, M., & Austin, M. (2013). Performance measurement challenges in nonprofit human service organizations. *Nonprofit and Voluntary Sector Quarterly, 43*(6), 1014-1032. doi:10.1177/0899764013508009

Cong, X., & Pandya, K. (2003). Issues of knowledge management in the public sector. *Electronic Journal of Knowledge Management, 1*(2), 25-32.

De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of big data based on its essential features. *Library Review, 65*(3), 122-135. doi:10.1108/LR-06-2015-0061

Ebrahim, A. (2002). Information struggles: The role of information in the reproduction of NGO-funder relationships. *Nonprofit and Voluntary Sector Quarterly, 31*(1), 84-114. doi:10.1177/0899764002311004

Farah, B. (2017). A value based big data maturity model. *Journal of Management Policy and Practice, 18*(1), 11-18.

Farkas, M. G., Hinchliffe, L. J., & Houk, A. H. (2015). Bridges and barriers: Factors influencing a culture of assessment in academic libraries. *College & Research Libraries, 76*(2), 150-169. doi:10.5860/crl.76.2.150

Gamage, P. (2016). New development: Leveraging 'big data' analytics in the public sector. *Public Money & Management, 36*(5), 385-390. doi:10.1080/09540962.2016.1194087

Gantz, J., & Reinsel, D. (2011). Extracting value from chaos. *IDC IView,* 1-12.

Gray, J. (1996). Data management: Past, present, and future. *IEEE Computer, 29*(10). http://arxiv.org/abs/cs/0701156

Gregory, A. G., & Howard, D. (2009). The nonprofit starvation cycle. *Stanford Social Innovation Review, 7(*4), 49. https://search.proquest.com/docview/217173506

Hawthorne, G., & Elliot, P. (2005). Imputing cross-sectional missing data: Comparisons of common techniques. *Australian and New Zealand Journal of Psychiatry,* 583-590.

Khatri, V., & Brown, C. (2010, Jan 1,). Designing data governance. *Communications of the ACM, 53*, 148-152. doi:10.1145/1629175.1629210 http://dl.acm.org/citation.cfm?id=1629210

Lakos, A., & Phipps, S. E. (2004). Creating a culture of assessment: A catalyst for organizational change. *Libraries and the Academy*, *4*(3), 345-361.

Lenczner, M., & Phillips, S. (2012). From stories to evidence: How mining data can promote innovation in the nonprofit sector. *Technology Innovation Management Review, 2(*7), 10-15. https://search.proquest.com/docview/1614473198

Levitin, A. V., & Redman, T. C. (1998). Data as a resource: Properties, implications, and prescriptions. *Sloan Management Review, 40*(1), 89. https://search.proquest.com/docview/1302987002

Luen, T., & Al-Hawamdeh, S. (2001). Knowledge management in the public sector: Principles and practices in police work. *Journal of Information Science, 27*(5), 311-318.

M+R. (2017). *2017 benchmarks study*. https://mrbenchmarks.com/

Malik, P. (2013). Governing big data: Principles and practices. *IBM Journal of Research and Development*, *57*(3/4), 1:13. doi:10.1147/JRD.2013.2241359

Mattia, A. (2011). An exploratory study on using stage theory to develop A data management measurement framework. *International Journal of Management & Information Systems (IJMIS), 15*(4), 123. doi:10.19030/ijmis.v15i4.5810

Mitchell, G. E., & Berlan, D. (2016). Evaluation and evaluative rigor in the nonprofit sector. *Nonprofit Management and Leadership, 27*(2), 237-250. doi:10.1002/nml.21236

Ndoye, A. (2010). Creating and sustaining a culture of assessment. *Planning for Higher Education, 38*(2), 28-39.

Nolan, R. L. (1979). Managing the crises in data processing. *Harvard Business Review*, *57*, 126. https://search.proquest.com/docview/227788117

Nolan, R. (1973). Managing the computer resource: A stage hypothesis. *Communications of the ACM*, *16*(7), 399-405. doi:10.1145/362280.362284

NTEN. (2012). *The state of nonprofit data*. https://www.nten.org/knowledge/reports-and-resources/

Redman, T. (1998). The impact of poor data quality on the typical enterprise. *Communications of the ACM, 41*, 79-82. doi:10.1145/269012.269025 http://dl.acm.org/citation.cfm?id=269025

Stoecker, R. (2007). The research practices and needs of non-profit organizations in an urban center. *Journal of Sociology & Social Welfare*, *34*(4), 97.

Thompson, N., Ravindran, R., & Nicosia, S. (2015). Government data does not mean data governance: Lessons learned from a public sector application audit. *Government Information Quarterly, 32*(3), 316-322. doi:10.1016/j.giq.2015.05.001

US Department of Health and Human Services [USDHHS]. (2012). *Navigating the family violence prevention and services program*. https://www.acf.hhs.gov/fysb/resource/fvpsa-state-guide

Weber, K., Otto, B., & Österle, H. (2009). One size does not fit all: A contingency approach to data governance. *Journal of Data and Information Quality (JDIQ), 1*(1), 1-27. doi:10.1145/1515693.1515696

Curriculum Vitae

Ashley M. Hernandez-Hall

4-H Assistant Professor WSU Extension

University of Nevada Las Vegas, PhD Candidate

Email: ahernandezhall@gmail.com

Education

In Progress Ph.D. Public Policy and Leadership, University of Nevada, Last Vegas, NV

2014 Master of Social Work, University of Nevada, Last Vegas, NV

2011 Bachelor of Social Work, University of Nevada, Last Vegas, NV

Employment

2019-present Assistant Professor, 4-H Faculty, Washington State University, Snohomish County

2018-2019 Director, SAFE Elementary Program, NyE Communities Coalition, Pahrump, NV

2017-2019 Graduate Research Assistant, University of Nevada, Las Vegas, NV

2016-2017 Graduate Assistant, University of Nevada Las Vegas, NV

2014-2016 Data Analyst, UNLV Libraries, University of Nevada, Las Vegas, NV

2011-2014 Social Work Graduate Assistant, University of Nevada, Las Vegas, NV

Honors And Awards

2018 Graduate Research Award, University of Nevada, Las Vegas

2017 Graduate Teaching Award, University of Nevada, Las Vegas

2015 Rising Star Award, National Association of Social Workers – Nevada Chapter

2011 Cum Laude Degree Honor, University of Nevada, Las Vegas

Publications And Creative Work

Peer-reviewed Journal Articles

Hoffman, S., & Hall, A. (2017). The Data Framework: A Collaborative Tool for Assessment at the UNLV Libraries. Journal of Electronic Resources Librarianship, 29(3), 159-167.

Creative Scholarship in Juried Events

Hall, A, & Hoffman, S. (2016). Using a Tool to Build a Culture of Assessment: The Data Framework, Library Assessment Conference, Arlington, VA, 2016. Washington D.C.: Association of Research Libraries.

Hoffman, S, & Hall, A. (2016). Visualizing Local Data: The Ithaka S + R Survey at UNLV, Library Assessment Conference, Arlington, VA, 2016. Washington D.C.: Association of Research Libraries.

Presentations

Hall, A. (2019, April). Nonprofit Data Management – A Stage Model. West Coast Nonprofit Data Conference, Phoenix, AR.

Hall, A. (2018, November). Tableau can Help: Using Data Visuals to Tell Your Story. American Evaluation Association Conference, Cleveland, OH.

Hall, A. (2017, March). Beyond Evaluation: Managing Nonprofit Data. American Society for Public Administration, Denver, CO.

Hall, A. (2016, November). Gateway Data: Visualizing Survey Data Helps Foster a Culture of Assessment at UNLV Libraries. Tableau Conference, Austin, TX.

Hall, A. (2016, September). Is Evaluation a Part of Your Practice? Fostering a Culture of Organizational Assessment in Human Services. NASW-NV Annual Conference, Las Vegas, NV.

Hall, A. (2015, September). Social Work in the 21st Century: Practicing efficiently and Ethically in Today's Digital Landscape. NASW-NV Annual Conference, Reno, NV.

Professional Service

University

2019-Present WSU Extension Island County 4-H Youth Development Coordinator Search Committee

Community

2019-present Glacier Peak Institute Board, WSU Representative, Board Member, Advisor

2018-2019 Pahrump Veterans Memorial Advisory Committee, Secretary

2017-2018 Pahrump Veterans Memorial Advisory Committee, Board Member

Review Activities

"Biosecurity: Public Speaking for Biosecurity Advocates 1: Creating a Persuasive Presentation" WSU Online Curriculum, Peer Review

Goodman, X., Perna, C., Juniel, P.M., & Weigel, R. (2016, January). Improving Capstone Papers for Baccalaureate Nursing Students: With an Evidence-Based Partnership Between a Health Sciences Librarian and a Nurse Educator. Presentation at In Tune with the Future through Vision, Visibility and Partnership Joint Meeting, Stanford University. Statistical Analysis Credit.

Juniel, P., Perna, C. A., Weigel, R., & Goodman, X. Y. (2016, May). An evidence-based approach to assessing writing of undergraduate nursing student's leadership capstone papers. Presented at the Medical Library Association Annual Conference and the International Clinical Librarian Annual Conference Medical Librarian Association and the International Clinical Librarian Annual Conference, Toronto, Canada. Statistical Analysis Credit.

Melilli, A., Mitola, R., & Hunsaker, A. (2016). Contributing to the library student employee experience: Perceptions of a student development program. Journal of Academic Librarianship, 42(4), 430-437. doi:10.1016/j.acalib.2016.04.005. Statistical Analysis Credit.

Community Engagement, Extension, And Outreach

2019

Conflict Management, 2-hour Volunteer Training Workshop, 22 participants, Presenter/Speaker, Coordinator/Organizer

Diversity and Inclusion, 1-hour Volunteer Training Workshop, 25 participants, Presenter/Speaker, Coordiantor/Organizer

Explore 4-H, 4-hour Community Outreach and Recruitment Event, 300 estimated participants, Presenter/Speaker, Co-Coordinator/Organizer

Professional and Scholarly Organization Affiliations

2014-2018	American Evaluation Association, Member
2012-2016	National Association of Social Workers – Nevada Chapter, Member
2012-2015	National Association of Social Workers – Nevada Chapter, Treasurer
2013-2014	University Association of Social Work Students, President
2012-2013	University Association of Social Work Students, Member
2013-2014	Phi Alpha Honor Society, President
2012-2013	Phi Alpha Honor Society, Member

PROFESSIONAL DEVELOPMENT

2019

Volunteer Engagement, Online and Everett, WA

4-H Program Days, Puyallup, WA

Connecting Military and 4-H Communities, Renton, WA

Discrimination, Sexual Harassment, and Sexual Misconduct Prevention, Online, WSU

Implicit Bias Awareness Training, Online, WSU

2017

R: Learning by Example, Boulder, CO

2016

Tableau Server, Las Vegas, NV

Tableau Desktop Advanced, Seattle, WA

2015

Tableau Desktop Fundamentals, Seattle, WA

2014

Microsoft Word, Essential Skills, Las Vegas, NV

Microsoft Excel, Formatting, Las Vegas, NV

Microsoft Word, Advanced Formatting and Editing, Las Vegas, NV

Microsoft Word, Graphics and Tables, Las Vegas, NV

Microsoft Word, Mail Merge, Las Vegas, NV

Microsoft Excel, Formulas and Functions, Las Vegas, NV

Microsoft Excel, Using Multiple Worksheets and Workbooks, Las Vegas, NV

Microsoft Excel, Charts and Graphs, Las Vegas, NV

Microsoft Excel, Data Tools, Las Vegas, NV

Microsoft Power Point, Essential Skills, Las Vegas, NV