MANAGING FOR PERFORMANCE: THE ROLE OF PROXIMITY IN PERFORMANCE MANAGEMENT

by

EUNJIN HWANG

(Under the Direction of Gene A. Brewer)

ABSTRACT

This study examines how public managers' proximity to the performance measurement process affects their perceptions of their agencies' performance measurement criteria, which is a form of measurement validity. The main variables of interest are managers' proximity to the performance measurement process and their accuracy in assessing the value content expressed in their agencies' performance measurement criteria. Inaccuracy or 'bias' is measured as the difference between managers' subjective perceptions of the values expressed in their agencies' performance measurement criteria, compared with objective measures derived from content analysis of the agencies' annual performance and accountability reports. The pertinent values are efficiency, customer service satisfaction, service quality, and social equity. The prevalence of each value is compared across time and in various stages of the policy process at the agency level. In this

study, the units of analysis are U.S. federal government managers in the 23 largest U.S. federal government agencies, enabling both individual and agency level analyses. Data come from the 2013 and 2017 annual performance and accountability reports for each agency (N=23), and the 2013 and 2017 Federal Managers Survey on Organizational Performance and Management Issues (N=2,762 in 2013 and N=3,114 in 2017). Multi-level models are employed in the study. The main results show that proximity affects managers' perceptions of their agencies' performance measurement criteria, and greater proximity can result in either overestimation or underestimation depending upon the form of proximity (subjective or objective). The study thus introduces the concept of proximity to the performance management literature and shows how it influences managers' perceptions of their agencies' performance measurement criteria.

INDEX WORDS: Performance management, Performance measurement, Measurement validity, Proximity, U.S. federal agencies, U.S. federal managers, Manager's involvement

MANAGING FOR PERFORMANCE: THE ROLE OF PROXIMITY IN PERFORMANCE MANAGEMENT

by

EUNJIN HWANG

B.A., Seoul Women's University, Korea, 2008

M.P.A.,Yonsei University, Korea, 2011

M.P.P., Arizona State University, 2014

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2020

MANAGING FOR PERFORMANCE: THE ROLE OF PROXIMITY IN PERFORMANCE MANAGEMENT

by

EUNJIN HWANG

Major Professor:	Gene A. Brewer
Committee:	Hal G. Rainey
	J. Edward Kellough

Electronic Version Approved:

Ron Walcott
Interim Dean of the Graduate School
The University of Georgia
August 2020

DEDICATION

This dissertation is dedicated to my father, Youngdon Hwang; to my mother, Sungwon Yang; to my younger sister, Hyejin Hwang; to my little brother, Woojin Hwang; to my brother-in-law, Minhyoung Lee; and to my lovely nephews Yejoo Lee and Haeseong Lee. Even when times were tough, my family supported me with a strong sense of accomplishment. I am incredibly fortunate to have you. Without your encouragement and support, it would not have been possible for me to complete this work to achieve my academic journey. My deepest appreciation and love to my family.

ACKNOWLEDGEMENTS

I would like to express my appreciation to the faculty who supported me throughout the dissertation process. I sincerely appreciate Dr. Gene A. Brewer, my major professor. He introduced me to the topic of performance and the idea of pursuing a career in research. He has always helped me since I first set foot in Baldwin Hall. He has always supported and trusted me, giving insightful advice, both personal and professional, when I have faced difficulties. This made everything possible. I would like to express my deep appreciation to Dr. Hal Rainey and Dr. Edward Kellough. They gave me valuable advice and feedback. I owe a great deal of my academic improvement to all of these committee members. I am also thankful to Dr. Seung-Yong Rho and Dr. Eun-Kook Lee, my mentors. Their tremendously helpful guidance and wisdom have been instrumental in making me a scholar.

I would like to express my special thanks to my cohorts Jordan Tobias Long, Samara Scheckler, Hadley Nobles, Dhanakorn Mulaphong, and Roger Qiyuan Jin. We share many good memories of our time spent in Baldwin Hall. Lastly, my family always gave me encouragement and support to continue whenever I felt exhausted in my pursuit of my doctorate degree in the U.S. My parents, my younger sister and her family, and my younger brother are my precious assets and serve as my safe harbor in my career.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

In the 1990s, government performance became a major concern in the public sector and a central theme in Public Administration theory and practice internationally (Brewer, 2000; Bouckaert & Halligan, 2007; Pollitt & Bouckaert, 2000; Talbot, 1999).[1] Ingraham (2005) states that "for much of the twentieth century—and certainly for the last 25 years—performance has been a siren's song for nations around the world" (p. 390). In the past decade or two, scholars have focused more intently on performance-related topics, including performance goals and objectives, performance measures, performance evaluation, and performance information use for managing organizational performance (Andrews, 2014). Scholars have generally agreed that the public sector should emphasize managing performance and developing and utilizing measures to improve satisfactory performance (Poister & Streib, 1999; Rendon, 2008; Sanger, 2008; Schick, 2001).

Performance is a contested concept that may mean different things to different people at different times and in different situations (Brewer & Walker, 2010; Hofer, 1983; Walker, Boyne, & Brewer, 2010). In the Public Administration field, scholars have attempted to delineate the concept of performance apart from political or policy objectives. They have defined the performance concept as a sequential set of activities that result in the achievement of goals.

[1] The term 'Public Administration' is capitalized when referring to the academic field of PA but written in lower-case when referring to public administration practice (Waldo, 1968).

Performance refers to "whether the agency does well that which it is supposed to do, whether people in the agency work hard and well, whether the actions and procedures of the agency and its members are well suited to achieving its mission, and whether the agency actually achieves its mission" (Rainey & Steinbauer, 1999, p. 13). Therefore, in many studies, satisfactory performance is seen as when the intended policy or program has achieved its expected outcomes (Giacchino & Kakabadse, 2003).

Research on Performance Measurement in the Public Sector

Performance measures have generally been understood as an impetus for organizational effectiveness that greatly exceeds their actual usefulness in practice (Poister & Streib, 1999). Therefore, performance measurement in contemporary government reforms has become "the hottest topic in government today" (Nyhan & Martin, 1999, p. 348). Most federal agencies are using a significant performance measures approach to track progress towards goals and make regular reports on their achievements and shortfalls (Aristiqueta, 1999; Poister & Steib, 1999; Radin, 2006). Even though a government agency may struggle to grasp the meaningfulness of performance measures (Ammons & Rivenbark, 2008), the enthusiasm of modernizing the public sector based on a performance-oriented management system remains high (Christensen & Laegreid, 2001; Pollitt & Bouckaert, 2000). While the validity of performance measures is a major issue in most systems, efforts to improve validity have sought to ensure that measures of performance are reliable, valid, and job-related.

The concept of performance in the public sector is a multidimensional construct (Fitzgerald & Moon, 1996; Otley, 1999) that includes diverse perspectives, such as quality, efficiency, effectiveness, responsiveness, and equity (e.g., Andrew, Boyne, & Walker, 2006; Van

Dooren, Bouckaert, & Halligan, 2010; Walker, Boyne, & Brewer, 2010). There are two difficulties in measuring performance in the public sector. First, public values and public performance are conceptually connected to each other and their relationship can be developed further (Van Dooren, Bouckaert, & Halligan, 2010). In the past, measuring efficiency was included as a standard criterion based on the scientific management approach. Recently, other perspectives, such as service quality and customer satisfaction, have been added (e.g., Poister & Streib, 1999; de Bruijn & Dicke, 2006; Holzer, Charbonneau, & Kim, 2009). Consequently, the current practices of public organizations in the United States tend to emphasize the quality of service, social equity, and citizen satisfaction (Holzer, Charbonneau, & Kim, 2009).

Another concern is the managerial role in performance measurement systems. Traditionally, managers need massive amounts of detail, including measures of workload, work procedure, and results, in order to improve program performance and align performance management systems (Franklin, 1999). This aspect of the managerial role is related to manager's accountability for results, and it involves both setting performance indicators and achieving performance goals (Aristigueta, Cooksy, & Nelson, 2001). Managers are, therefore, central actors caught in the crossfire between performance measurement and performance management. Despite their important role in the organization, many studies report that managers' perceptions are likely to differ significantly depending on their purpose, intentions, and contribution to the performance measurement process. They are often obliged to justify organizational activities and produce performance reports. In addition, managers' knowledge, their attitudes toward work, job satisfaction, and perception of organizational performance unintentionally produce positive and optimistic responses (e.g., Brewer, 2005). Since their perceptions could be overly subjective, their assessments of organizational performance can be overestimated, resulting in bias that can

decrease performance measurement validity. Furthermore, managers are not always well-informed of the complexities of government work, which can have high levels of ambiguity and conflict, and then may adopt measures that are too weak or too demanding (Bouckaert & Peters, 2002).

Performance Management and Measurement Validity

Performance management involves both organizational aspirations and performance achievements. Organizational aspirations refer to desired performance levels in organizational outcomes (Shinkle, 2012). Organizational aspirations have been called goals, reference points, and targets based on goal-seeking and analytical planning behavior (Ansoff, 1979, 1987). Therefore, organizational aspirations play a significant role in strategic decision making because setting goals and objectives is a component of scientific management (Taylor, 1911) and strategic planning (e.g., Chandler, 1962; Hofer & Schendel, 1978). Managers set organizational goals to correspond with their strategic direction (Fiegenbaum et al., 1996), enhance performance (Carver & Scheier, 1981; Eisenhardt, 1985; Hamel & Prahalad, 1989), and provide measures of success (Kaplan & Norton, 2001; Meyer, 2002). Managers establish organizational goals and objectives based on past aspirations, performance, historical traditions such as values and inherited attributes, inertia, organizational capabilities, external stakeholders, and "managerial perceptions of need" (Ansoff, 1987, p. 511). Therefore, managers' perceptions of organizational goals are based upon their cognition (Kiesler & Sproull, 1982) and they "select explicitly or implicitly the aspirations for an organization" (Shinkle, 2012, p. 423). Since an agency's annual performance and accountability report (e.g., the agency plan) reflects the managers' perceptions of those needs and the centrally established goals of the agency, there is,

therefore, an implied underlying congruence between an agency's goals and the managers' perception of those goals.

It is important to examine whether the agency's annual report and the managers' perceptions of the organizational goals are congruent. Measurement coherence or congruence can be achieved when managers correctly recognize that their agency's goals, or the agency's annual reports, accurately capture the managers' aspirations (Mitchell et al., 2013). The lack of coherence implies gaps between measures taken from the agency's annual reports and the managers' perceptions of goals stated in the agency's annual reports. That is, low congruence can be thought of as poor measurement validity, delivering challenges to many organizations (Mitchell et al., 2013). On the other hand, coherence between the agency's annual report and the managers' perceptions indicates measurement accuracy and high validity. Thus, in this study, performance measurement validity includes measurement coherence and congruence between value content in agencies' annual reports and managers' perceptions of those values.

This study uses subjective measures for analyzing managers' perceptions of agency goals and reflected values, and archival-based measures for analyzing performance indicators in the agency's annual reports. This study then seeks to examine whether managers accurately perceived their agency's goals that were expressed in agency plans. In addition, this study examines the coherence, congruence, or measurement validity of those subjective and objective measures. Incongruence, or managers' bias, refers to potential gaps between subjective and objective measures – between managers' perceptions of the agencies' goals and archival-based measures taken from the agency's annual reports.

Comparison of Subjective and Objective Measures

As stated above, this study uses subjective measures for analyzing managers' perceptions of agency goals and reflected values, and archival-based measures for analyzing performance indicators in the agency's annual reports. Therefore, this study develops subjective measures from the survey data and objective measures from agency annual performance and accountability reports, which is an archival-based source.

A crucial factor for this analysis is the availability of information. Surveys of Federal Managers on Organizational Performance and Management Issues (FMOPMI) is a valuable source of data and it is used to analyze managers' perceptions of their agency's performance measurement process. In this survey, organizational goals and corresponding public values are analyzed via managers' subjective assessments. On the other hand, managers establish organizational goals and objectives based on past aspirations and managerial perceptions of need, and these perceptions are also reflected in the Performance and Accountability Reports (PARs). Thus, it works both ways: performance measurement is influenced by managerial and policy decisions, and managers intentionally select and integrate performance measures into key management systems or decision processes (Clay & Bass, 2002; Poister & Streib, 1999). PARs are archival-based data, which is more objective; however, they still contain managerial perceptions of need based upon organizational capabilities and past aspirations. Put another way, the assumption is that agencies and their managers will embrace a common set of performance objectives that emphasize chosen values. This assumption can be thought of as value congruence.

Even though there are differences in the ways data are collected, there is an underlying presumption that subjective assessments of the public values reflected in performance goals should correspond closely with objective data (e.g., Forth & McNabb, 2008; Wall, Michie,

Patterson, Wood, Sheehan, Clegg, & West, 2004). To date, however, few empirical studies have analyzed performance measures by comparing objective and subjective types of data, partly because there are few suitable data sets with matching measures (Klimoski & Inks, 1990; Mero, Guidice, & Anna, 2006). Therefore, comparing subjective and objective measures is clearly important because it provides evidence of the extent to which managers' perceptions of performance criteria are congruent with agency PARs.

How the Policy Process Model Relates to Agency Performance and Accountability Reports

A policy process is a comprehensive approach for understanding performance dimensions (Boyne, 2002) because the policy process model includes different stages (e.g., input-process-output-outcome), and diverse organizational goals reflected different public values are contained in each stage. Each stage has different characteristics. For example, process and output stages are more specific than outcomes, and outcomes are broader and vaguer than process and output (Kelly, Muers, & Mulgan, 2002).

In addition, since the policy process has different characteristics of stages, the policy process could be applied to the agency's performance measurement diagram in the Agency Performance and Accountability Reports. A federal agency follows agency strategic planning section 230 and all agencies have strategic goals, objectives, and performance indicators in their annual reports. Strategic goals are broader long-term statements that reflect the agency's mission while performance indicators are specified activities that are measurable, aiming at distinct targets. In this perspective, each stage of the policy process model that has different characteristics that could be related to the strategic goals, objectives, and performance indicators

of the performance measurement diagram in the agency's annual report. This will be further discussed in Chapter 2.

Purpose and Significance of the Dissertation

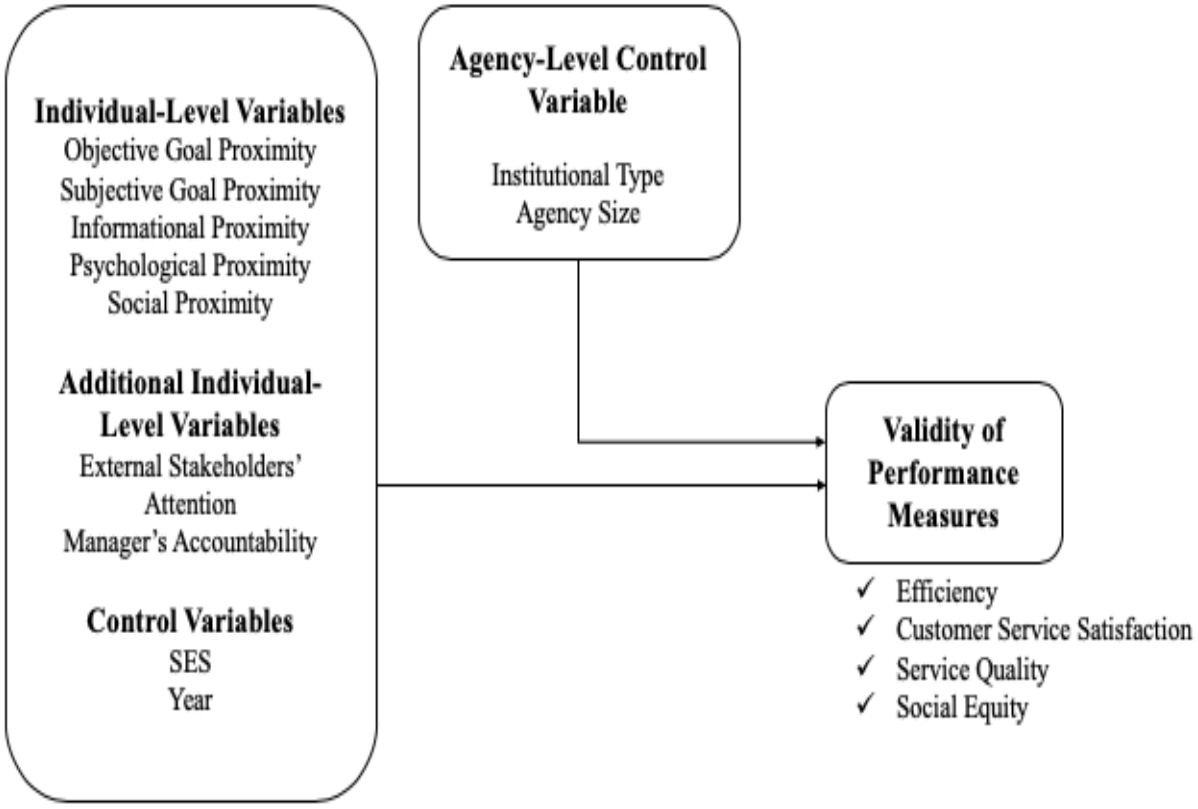
The previous literature has reached a consensus that the characteristics of performance measurement should emphasize validity, reliability and relevance, providing an analysis of how well program results meet expected goals (Franklin, 1999; Radin, 1998; Tigue & Strachota, 1994; Wholey, Hatry, & Newcomer, 2010). Performance measurement is therefore an important concept in the public management literature. The growing number of studies on performance measurement has focused on the questions of what conditions affect the validity of performance assessment and how measurement validity can be achieved (Brewer, 2006; Favero & Bullock, 2014; Jung, 2013; Meier & O'Toole, 2013a). Recently, scholars have noted that performance measures could be overly optimistic and correlates poorly with archival performance measures (e.g., Meier & O'Toole, 2013b, Favero, Anderson, Meier, O'Toole, & Winter, 2015).

Despite chronic concerns regarding performance measurement, little empirical research has attempted to examine measurement validity. Of the 111 articles published by the *Public Performance & Management Review* from 2017 to 2019, only 16 articles examined performance measurement validity. Only 10 of these 16 articles employed both objective and subjective data. Since *Public Performance & Management Review* is the journal that mainly focuses on performance-related topics in the public sector, other journals probably have similar or less frequency of objective data use in performance measurement articles. Therefore, this study will explore both subjective and archival-based data and compare these two sources of data in a set of public organizations to help gauge perceptual errors and biases (Starbuck & Mezias, 1996).

Therefore, based on objective and archival-based data, this study tracks trends of the value content of agency performance measurement criteria and compares them across agencies. There is potential disconnect between managers' perceptual data of value content and archival based data. For example, some agencies may not distill their missions into a truly accurate agency plan. Managers might be more correct about their agency's values in the real world. Yet the gap still exists. Therefore, this study still could contribute to the performance measurement validity.

In addition, since managers play a key role in performance-related activities, proximity effects will be studied. Proximity has been examined in other social science fields (Knoben & Oerlemans, 2006). In Public Administration, however, few studies have examined the relationship between proximity and performance, and previous studies have only focused on geographical closeness as their form of proximity. The proximity effects developed in this study can be more broadly applied to performance management study. Taken together, the goal of this study, therefore, is to examine performance measurement validity by assessing the gap between archival-based and perceived measures of U.S. government federal agencies. Furthermore, this study investigates the main forms of proximity that affect performance measurement validity as shown in Figure 1.1.

Figure 1.1 Research Framework



This study aims to make several contributions to the field of Public Administration and management. First, it will track trends of the value content of agency performance measurement criteria and compare them across agencies. Since the value content of performance measurement criteria unfolds in various stages of the performance management process, tracking changes in value content expressed in the agency's report across the stages of performance is meaningful because understanding what value is expressed in each stage reveals some patterns of performance measurement validity across time. This work will show whether the government agencies have an imbalance in value content based on their mission, which can help managerial decision making when planning activities to achieve the agency's mission. Second, the introduction of the concept of proximity into the study of Public Administration can help

organizations understand their performance measures better. The concept of proximity has received little attention in the field of Public Administration, partly due to its conceptual ambiguity. In Public Administration, some studies have employed the proximity, which refers to their physical closeness. Yet, distance in physical terms has been regarded as an incomplete view of how people experience proximity (Wilson et al., 2008). The work undertaken here will hopefully illustrate new ways to study performance measurement validity of value content and proximity effects in public organizations. This work aims to provide insights that will help strengthen performance measurement and management systems in government. Lastly, as defined in the earlier paragraph, this study examines performance measurement validity based on measurement coherence and congruence between value content in the agency's annual reports and managers' perceptions of value content.

The study will also estimate the gap between employee perceptions of value content in their agency's performance measures and archival-based measures taken from the agency's annual performance and accountability report. The expectation is that a significant gap exists between managers' perceptions of the values reflected in performance measurement criteria and independent assessments based on the agency's annual performance and accountability report. Estimating the measurement validity will provide useful information on the degree to which federal agency performance measures are distorted in the eyes of federal agency managers and establish what part of this distortion derives from their proximity or distance from the performance management process. This is actionable information for policymakers and public managers overseeing the performance management process. It could, for example, partially correct the distorted instrumentation that plagues perceptual measures, which are otherwise easier to collect and more encompassing than most purported objective measures. Other benefits

might accrue for researchers who are trying to develop more valid, reliable, and complete measures of performance in the public sector.

Main Research Questions

Recognizing the necessity of studying performance measurement validity in the field of Public Administration, this dissertation intends to answer the following research questions:

1. Are there significant differences in performance measurement validity in value content when comparing measures from 2013 and 2017?
2. Is there a gap between individual perceptions of performance measurement validity (taken from Surveys of Federal Managers on Organizational Performance and Management Issues) and archival-based measures (taken from Performance and Accountability Reports), and if so, does proximity help explain this gap?
3. What other factors affect the gap between archival-based and perceptual measures of performance measurement validity?

Dissertation Structure and Outline

This dissertation is organized into six chapters: Chapter 1, the introduction, provides a brief background for the concept and previous research on performance measurement. The purpose of this study is explained, and an outline of the dissertation is provided. Chapter 2 presents definitions of key terms and the central term of performance measurement, describes its characteristics, and discusses performance measurement in practice based on federal agency annual performance and accountability reports, which enable comparison between manager perceptions and archival-based measures reflecting the reality of performance measurement.

Chapter 3 explores potential factors that may affect performance measurement validity. Diverse forms of proximity effects as main predictors of managers' perceptions on performance measurement are explored. To address the research questions proposed, several hypotheses are formulated throughout the chapter. Chapter 4 presents the research design and describes the data and methods. Chapter 5 reports on the statistical relationships between manager proximity and performance measurement validity in 23 U.S. federal government agencies. Finally, in Chapter 6, the results are summarized and some suggestions for future research are offered.

CHAPTER 2

PERFORMANCE MANAGEMENT AND PERFORMANCE MEASUREMENT

Many scholars have devoted considerable attention to exploring performance management and measurement in the public sector. Both are tools to help organizations improve their effectiveness and strengthen policy processes, not only by providing useful information for planning purposes and resource-allocation decisions, but also by increasing accountability and transparency in government programs and instilling a culture of 'managing for results' (Kamensky & Abramson, 2002; Moynihan, 2008; Newcomer, 2007; Williams, 2003).

Several publications focus on linking performance measurement and management (e.g., de Lancer Julnes, 2008; Moynihan, 2008). In order to understand this linkage, Bouckaert and Halligan (2007) describe three stages of an intellectual evolution towards managing for results: performance administration, management of performances, and performance management. At the first stage, 'performance administration,' performance measurement exists in a traditional setting, such as Taylor's scientific management and Weber's description of bureaucracy. In this stage, it predominantly focuses on input and output, concentrating on a result-oriented vision of performance (Fry, 1989). Performance administration, therefore, emphasizes efficiency (Bouckaert & Halligan, 2007). The second stage of performance evolution is 'management of performances,' a shift from the first stage. Performance measures tend to emphasize resources and products/services delivered. This stage includes three dimensions: measuring performance using a pre-designed monitoring system for improving validity, reliability, and quality (such as a

balanced scorecard); incorporating specific management functions (such as personnel, strategy, and communication); and linking performance information to performance. The third stage is called 'performance management,' which spans from input through outcome. It is concerned with allocating responsibilities for performance. All three stages of an intellectual evolution focus on performance measurement for performance management by emphasizing different perspectives, such as the result-oriented scheme at the first stage, building systems dedicated to systematic measurement in the second stage, and shifting to an accountability focus in the third stage. Even though the focal points of each stage are different and changing, the importance of managing performance is not. Reviewing the linkage between performance measurement and management, this study focuses on performance measurement validity based on the congruence and coherence of agency-level measures taken from annual reports and managers' perceptions, as mentioned in Chapter 1.

This chapter further reviews performance management and performance measurement based upon empirical studies and related practitioner experience reported in the literature. The overview begins by exploring the definition and the historical development of the performance management system of the U.S. government. The review then describes the concept of performance measurement and discusses how the current performance measurement system in the U.S. government developed. Then, several challenges of performance measurement in the public sector are examined and the concept of 'public value,' which is relevant to performance management and measurement challenges, is introduced. These discussions culminate in a set of hypotheses.

Performance Management

Definition of Performance Management

The concept of management is "the process of making an organization perform well through defining organizational goals and making decisions about the efficient and effective use of organizational resources" (Anderson, 1988, p. 8). Performance management can be described as the strategies for improvement of an organization's performance with managers' and employees' attention (Andrew, 2014). Performance management includes allocating resources, setting and aligning goals, measuring performance, monitoring strategies, employing a reward system, and collecting relevant data (Behn, 2003; Poister & Streib, 1999; Yang & Hsieh, 2007). In addition, performance management includes managing human resources by motivating performance, which includes "helping employees develop skills, building a performance culture, determining who should be promoted, eliminating poor performers, and helping to implement business strategies" (Lawler, 2003, p. 1).

In the era before performance management became relatively common in the public sector, Niskanen (1971) contends that public organizations were less efficient than companies in the private sector due to the monopoly environment, and duplication and overlap of public service production with rent-seeking pressures for bureau expansion (Miller & Moe, 1983). Jones and Kettl (2003) argue the following:

> Governments are inefficient, ineffective, too large, too costly, overly bureaucratic, overburdened by unnecessary rules, unresponsive to public wants and needs, secretive, undemocratic, invasive into the private rights of citizens, self-serving, and failing in the provision of either the quantity or quality of services deserved by the taxpaying public. (p. 1)

These criticisms have led to changes in the way government is managed. Government aspires to improve performance by implementing reforms and new management practices.

Historical Development of Performance Management of the U.S. Government

U.S. government agencies have sought proper performance management systems throughout history. Mosher (1981) breaks down American administrative history and portrays the twentieth century as a period dominated by 'government by the efficient' and 'government by managers.' The era of government by the efficient treats administration as a science led by the goal of efficiency based on principal-agent theory in the public sector, and emphasizes centralized, top-down regimes, a stable political environment, and one-size-fits-all policies (Dubnick & Frederickson, 2009; Radin, 2006). The first evidence of the development of a formal performance management system in public administration emerged during the Progressive era when Woodrow Wilson (1887) and others advocated for a scientific approach to the field, emphasizing the importance of efficiency, neutral competence, and a professional orientation. The Civil Service Reform Act (known as the Pendleton Act) was enacted in 1883 and established the U.S. Civil Service Commission. The New York Bureau of Municipal Research (NYBMR) was established in the early 1900s. The NYBMR adopted scientific techniques, linking resources to governmental objectives, municipal statistics, and cost accounting (Williams, 2002). They believed that "wasteful, ineffective government could not serve democracy well" (Dahlberg, 1966, p. 31). In this era, when specifying targets, program administrators tried to name easily achievable outputs and outcomes rather than more difficult ones. Measuring workload and workers' efficiency were clearly part of the scientific

management approach at the turn of the century (Ridley & Simon, 1943; Taylor, 1911). Williams (2003) illustrated that pressure for performance measurement existed in the Progressive era as reformers sought to decrease corruption and inefficiency by introducing tools that would make managers more accountable.

A second generation of performance management followed World War II (Ravindran, 2007; Yewlett, 2017). Scholars in this era emphasized the rational techniques of cost-benefit analysis and public policy choice, resulting in the development of management tools such as Planning, Programming and Budgeting System (PPBS), Management by Objectives (MBO), and Zero-Based Budgeting (ZBB) (Bouckaert & Halligan, 2007). The Hoover Commission (1949, 1955) encouraged employing performance-based budgets to reduce expenses. In 1962, the Bureau of the Budget started the productivity project with an annual report for the President and Congress (Kull, 1978). The Nixon administration also established a national productivity commission. The Grace Commission under the Reagan administration in the 1980s focused on efficiencies, emphasizing cost savings and minimal government. On a larger scale, the U.S. government has developed performance management systems for federal agencies based on job-related performance standards under the Civil Service Reform Act of 1978 (Frederickson, 1980), and the strategic planning approach that aimed to increase agencies' capacity and productivity to help public managers make better decisions (Balaboniene & Vecerskiene, 2015; Dusenbury, 2000; Rainey, 2014). The emphasis on efficiency was then resurrected by New Public Management (NPM), which was widely influential in the 1990s (Dunleavy & Hood, 1994; Hood, 1991). Osborne and Gaebler's 'reinventing government' program in 1992 suggested following best practices and instilling an entrepreneurial spirit in the public sector. This movement embraced the market principles of the private sector and led to distinctive changes in public

administration and public management by emphasizing performance measurement systems and competition between service providers (Dunleavy & Hood, 1994).

The clearest manifestation of performance management in the U.S. was the Government Performance and Results Act (GPRA), which was passed in 1993. According to GPRA, federal agencies were mandatorily and intensively involved in performance management activities, including setting program goals and measuring results (Piotrowski & Rosenbloom, 2002). GPRA required all federal agencies to create a strategic plan with goals and objectives that would be quantifiable and to specify objective performance indicators for their anticipated achievements during the fiscal year (Berman, 1998; Kettl, 1994; Rainey, 2014). According to Public Law 103-62, which enacted this initiative, the purposes of GPRA are to:

> (1) improve the confidence of the American people in the capability of the federal government, by systematically holding federal agencies accountable for achieving program results; (2) improve program performance by requiring agencies to set goals, measure performance against those goals and report publicly on progress; (3) improve federal program effectiveness and public accountability by promoting a focus on results, service quality and customer satisfaction; (4) help federal managers improve service delivery by requiring that they plan for meeting program goals and by providing them with information about program results and service quality; (5) improve congressional decision-making by providing more information on achieving statutory objectives and on the relative effectiveness and efficiency of Federal programs and spending; (6) improve internal management of the federal government; and (7) improve usefulness of

performance and program information by modernizing public reporting" (Government Performance and Results Act of 1993, p. 285).

The Government Accountability Office (GAO), the investigative arm of Congress, has compiled and reported on the successes and challenges of federal agencies in the performance arena (Julnes, 2006; Newcomer, 2007). Moreover, the Government Performance Project (GPP), which evaluated the capacity of management systems in government entities, was prominent in the twenty-first century and focused on five management areas: financial management, human resources management, capital management, information technology management, and managing for results (Rainey, 2014). The George W. Bush administration tried to clarify each agency's mission and program goals for improving their performance through the Program Assessment Rating Tool (PART) created by the U.S. Office of Management and Budget to rate all federal programs on their effectiveness. The PART consisted of about 30 questions, and managers needed to report the effectiveness of their program based on evidence, such as program purpose and design, program management, and program results (Newcomer, 2007). Many public organizations across different levels of government had successfully implemented the scorecard approach, linking tangible strategies with performance measurement (Holmes, Amin Gutiérrez de Piñeres, & Douglas Kiel, 2006; Wilson et al., 2014; Kaplan & Norton, 2001). The PART process was explicit and systematic, but some critics complained that since the unit of analysis for PART was government programs, there was a disconnect between broad agency performance goals in GPRA performance reports and the specific goals of certain programs (Newcomer, 2007).

In addition, the GPRA Modernization Act of 2010 (GPRAMA) was designed to further strengthen program performance by prioritizing goals, assigning more specific responsibilities to agency officials, and reviewing progress quarterly (GAO, 2011). It developed a balanced set of performance measures, which reinforced the need for agencies to have a variety of measures. Charbonneau and Riccucci (2008) have argued the need for social equity measures in performance measurement because performance indicators have overemphasized efficiency measures. Based on the results of comparison of performance measures in policing, they concluded that social equity indicators should "be part of a balanced mix of performance measures for government services" (Charbonneau & Riccucci, 2008, p. 616). Recently, the President's Management Agenda (PMA) further emphasized the use of performance measures and other 'evidence' to justify the continued existence and funding of all federal programs (The Performance Institute, 2018). One of the major themes under the Trump administration is that each program should identify critical performance metrics and show constant improvement, shifting government activities from low-value to high-value work to "improve the ability of agencies to deliver mission outcomes and provide excellent service" (p. 1). Moreover, the Program Management Improvement Accountability Act (PMIAA) was enacted in 2016. This law aimed to improve program and project management practices in the federal government, and it requires government wide standards for program management (Mulvaney, 2018). The federal government conducts portfolio reviews annually, establishing a five-year strategic plan for program management. Regardless of each administration's characteristics and policies, and the disconnects between their rhetoric and practice, they have devoted growing attention to improving government performance, and they have increasingly focused more attention on performance measurement and related topics.

The Logic of Performance Management

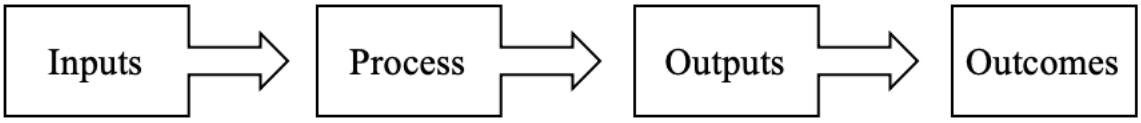
Conceptual frameworks for organizational performance in government include the 3Es (economy, efficiency and effectiveness) and the inputs-outputs-outcomes (IOO) model / inputs-process-outputs-outcomes (IPOO) model, both of which conform to the policy process model (e.g., Boyne, 2002; Talbot, 1999). The term economy, which is the most straightforward element of performance, is defined as "the cost of procuring specific service inputs of given quality" while the concept of efficiency refers to "the cost per unit of output" and effectiveness denotes "the achievement of the formal objectives of services" (Boyne, 2002, p. 17). Efficiency and effectiveness are considered a significant part of government performance indicators. All three elements of the 3Es model may apply to all policies in the public sector. It may also be necessary to consider other 'E's' in public administration, such as equity, environment, and ethics (Otrusinova & Pastuszkova, 2012).

This broader set of criteria is covered by the policy process model (input-process-output-outcome) for performance that is the currency of performance management as shown in Figure 2.1 (Talbot, 1999). The policy process model provides an explicit and comprehensive approach for understanding performance dimensions (Boyne, 2002; Sorber, 1993; Talbot, 1999). This sequential model has been broadly adopted in practice for decades and it is a systematic framework for an effective performance management system not only in the public sector but also in the private (Behn, 2003; Kaplan & Garrett, 2005). Hatry (2006) provides the logic of the public performance model which describes a sequence of inputs (e.g., resources, investments), activities (e.g., services, processes, strategies, methods), outputs (e.g., tangible products delivered by a program), and end outcomes (e.g., expected changes and results). Inputs include

staffing, workload, equipment, bureaucratic support, and spending while processes are the activities and progressions toward agency goals containing human resources processes, organizationally specific processes, information technology, and budgeting (Alach, 2016; Wholey, 1999). Outputs are defined as the goods and services produced by an organization including the quantity of a service and its quality to the public or others (Alach, 2016; Boyne, 2002). Outputs help track a program's progression for achieving its outcome. Outcome, which is important and meaningful to the public generally, refers to the intended and ultimate results, such as customer satisfaction, high service quality, and organizational change (Wholey, 1999). In a logical connection between outputs and outcomes, performance measures should distinguish between them, with outputs supporting outcomes because their relationship is more complicated in the public sector than in private services (Smith, 1996).

There are several advantages to using the policy process model (input-process-output-outcome). One of the fundamental purposes of public organizations is to distribute resources with social outcomes, while private sector organizations track profitability and mainly measure economic value. Therefore, performance indicators of public organizations based on the logic of the public performance model include not only cost-effectiveness or value for money rhetoric but also social value. Second, since each stage in a logical connection is linked closely to the other stages, performance indicators provide information about the relationship between contributions including costs, process, and consequences (Boyne, 2002). Through the policy process model (input-process-output-outcome), this dissertation examines how process, output, and outcomes measures have changed over time in government agencies and discusses the similarities and differences of performance measures across government agencies.

Figure 2.1 Public Performance Model



Source: Adapted from "Public performance—towards a new model?" by C. Talbot, 1999. *Public Policy and Administration*, 14(3), 17.

Performance Measurement

Definition of Performance Measurement

Performance measurement in the public sector is complex. Government programs and activities are difficult to measure due to differences in objectives, lack of quantitative measures for the values of outputs, and organizational characteristics (Brown & Pyers, 1988). Even though it is challenging, defining performance measurement in the public sector helps to clarify the concept and its purposes. Performance measurement refers to "quantifying, either quantitatively or qualitatively, the input, output, or level of activity of an event or process" (Fryer, Antony, & Ogden, 2009, p.393). Hatry (2006) also describes performance measurement as a process used in order to make accurate diagnoses for a path of progress, identifying the activities required. GAO (2011) further states that "performance measures may address the type or level of program activities conducted (i.e., the process), the direct products and services delivered by a program (i.e., the outputs), or the results of those products and services (i.e., the outcomes)" (p. 2). According to these descriptions of performance measurement stages, managers can collect and track information of an agency or specific program's inputs, activities, outputs, and outcomes, helping the agency or program clarify desired outcomes and the level of its attainment (Wholey, 1999; Wray & Hauer, 1997).

Many scholars suggest that the fundamental purpose of performance measurement is to "track selected performance measures at regular time intervals so as to assess performance and

enhance programmatic or organization decision making, performance, and accountability" (Poister, 2008, p. 15). Also, the use of performance measurement helps to justify resource requirements and improve employees' motivation (Ammons, 1995; Behn, 2003; Poister & Streib, 1999). According to Behn (2003, p. 593), public managers have purposes that are specific and distinctive and they use performance measurement to "(1) evaluate; (2) control; (3) budget; (4) motivate; (5) promote; (6) celebrate; (7) learn; and (8) improve." The general functions of performance measurement as an explicit tool are also demonstrated by Sorber (1993, p. 63): "(1) providing early warning on the developments of output and outcome; (2) improving allocation of resources; (3) improving the efficiency and effectiveness of production; and (4) improving agencies' or manager's accountability." Additionally, Hatry (1999, p. 157-158) suggests that performance measurement is used to carry out the following tasks: "(1) respond to elected officials' and the public's demands for accountability; (2) make budget requests; (3) do internal budgeting; (4) trigger in-depth examinations of performance problems and possible corrections; (5) motivate; (6) contract; (7) evaluate; (8) support strategic planning; (9) communicate better with the public to build public trust; and (10) improve." Given the functions of performance measurement according to the aforementioned scholars, we can conclude that performance measurement is not an end in itself, but "is intended as a means" for effective performance management (Julnes & Holzer, 2001, p. 694). As a result, scholars have wondered whether management matters for performance, and if so, what elements of management matter most (Boyne & Walker, 2005; Walker, Brewer, & Boyne, 2010).

Performance Measurement Systems in the U.S. Government

Poister (2008) states that a significant change has occurred in the current generation of performance measurement systems in the public sector. He described the change as follows:

> In contrast to earlier attempts at developing performance measurement systems, which often appeared to be less purposeful, less focused, and less well aligned with other evaluative and decision-making processes, the current generation of measurement systems are more mission driven and results oriented. Often, this new performance measurement approach is tied to a strategic plan, emphasizes the customer perspective, measures performance against goals and targets, and incorporates measurement systems in other management processes in meaningful ways. (p. 9)

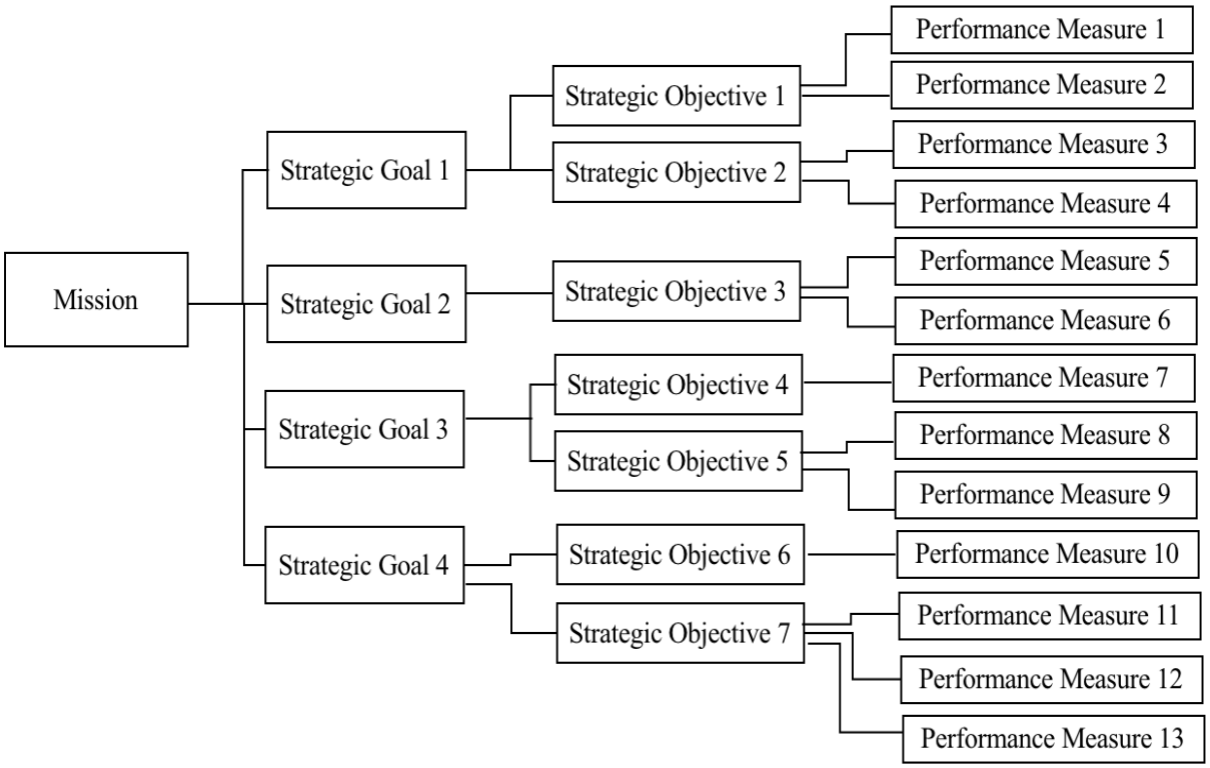
Based on the new performance measurement approach, by February each year, U.S. federal government agencies report performance goals, measures, and progress towards meeting those goals in their strategic plans to inform Congress and the American people about what they expect to accomplish on their behalf (GAO, 1992; WhiteHouse.gov). In the annual performance and accountability report, each agency first establishes a strategic plan that guides planning for program implementation. The agency's report and strategic plan develop a comprehensive mission and long-term vision statement that the agency aims to accomplish by stating policy intentions that cover the core functions of the agency (Dusenbury, 2000; Wholey & Hatry, 1992). This is meant to help agencies in several ways: an agency can identify and clarify its missions and goals as purposes and directions; and the agency can improve its ability to anticipate the future and make decisions by identifying issues, opportunities, and problems. Therefore, an agency's report and strategic plan, which is the foundation of the agency's planning system to implement the strategies needed to achieve goals, provide the context for

decisions about performance goals, priorities, strategic human capital planning, and budget planning for improving program effectiveness and accountability. Federal agencies should provide the details published in their Annual Performance Reports and on the Performance.gov website. Through a review of an agency's report and strategic plan, agency performance is analyzed and monitored, and strategic goals and objectives are revised accordingly. The agency's report and strategic plan outline its progress with specific challenges, opportunities, and progress based on the agency's mission, providing the goals' priority levels, agency capabilities, and the work environment. Therefore, an agency's annual report and strategic plan are a deliberative tool, not only providing the agency with guidance for planning program implementation, but also formulating an opportunity to engage external entities, such as Congress, the public, and the agency's stakeholders. External parties may also provide their ideas, expertise, criticisms, and assistance to accelerate progress.

As an agency's strategic plan links objectives to programs, the plan also defines missions, strategic goals, strategic objectives, and key performance indicators whereby the agency holds itself accountable (Dusenbury, 2000). Strategic goals, objectives, and performance indicators are described by agency strategic planning section 230, as illustrated in Figure 2.2. *Strategic goals* (SGs) are identified in both the agency's report and strategic plan, and normally comprise the broad and long-term statements that reflect the agency's mission in order to address needs, challenges, and opportunities. Each strategic goal should be supported by a suite of 2-10 mission-oriented strategic objectives and performance goals. *Strategic objectives* (SOs) include clear targets and directions for specific activity toward achieving the agency's long-term mission. Strategic objectives involve more specific performance goals and indicators related to the agency's mission (performance.gov). Strategic objectives are directly linked to agency

strategic goals, emphasizing the outputs of agency activities. When setting and developing each strategic objective, the agency should contemplate how to measure agency actions and mark progress toward achieving objectives through a set of performance indicators. *Performance indicators* (PIs) are quantifiable, measurable, and time-based statements of intent. They gauge the actual impact on specified activities and results under guidance from the Office of Management and Budget (OMB). Performance indicators are specific and distinct targets that are appropriately based on budgeting and resource allocations (Cassidy & Kendis, 2011).

Figure 2.2. Performance Measurement Diagram



Source: Adapted from "Performance measurement: an index approach for federal agencies" by J. Cassidy and S. Kendis, 2011. Paper presented at the PMI® Global Congress 2011, North America, Dallas, TX.

Factors Influencing Performance Management and Measurement

Many factors influence performance measurement in the public sector, including manager's accountability and external stakeholder attention. Bozeman and Straussman (1990) elucidate that *manager's accountability* is an important resource to achieve successful performance in an organization (e.g., Hatry, 1997; Ossege, 2012). Since accountability affects public managers' work behavior, recognizing the importance of accountability is necessary in public organizations. There are two types of accountability: process and outcome (Patil, Vieider & Tetlock, 2014). This study focuses on outcome-based accountability which is highly linked with personal motivation (Ryan & Deci, 2006) and many studies have confirmed that motivation moderates the degree to which a manager's behavior is affected by accountability (e.g., Furiady & Kurnia, 2015; Ossege, 2012; Ryan & Deci, 2006). According to the previous research, managers are more likely "achieving desirable societal outcomes" (Bovens, Schillemans, & Hart, 2008, p. 232) and may intentionally adjust performance grades because they are accountable for impacts that are deemed most important to them over outputs and process (Klimoski & Inks, 1990; Mero, Guidice, & Anna, 2006). This link between managers' behavior and performance suggests that accountability stimulates public managers to engage in tasks in order to perform well (Ossege, 2012).

Stakeholders play a key role in achieving the objectives of organizations, including businesses, media, employee unions, and environmental organizations (Anderson, Brewer, & Leisink, in press). Especially, external stakeholder pressure from elected officials and citizens is an important factor for managing performance in the public sector (Wang & Giannakis, 1999). Some studies have found that external stakeholders exert positive effects on performance. For example, the Committee on Governmental Affairs (1993) reports that Congress has an oversight

role that provides an influential tool for strengthening governmental management (Moynihan & Lavertu, 2012). On the other hand, some studies have found that external stakeholder attention has a negative effect on performance. For example, as reported in the National Performance Review (NPR) issued in 1993, external stakeholders introduce several obstacles that restrict improving performance. One of these obstacles is from the Inspector General's 'heavy-handed' approach. Managers should follow every rule and document every decision. They are even discouraged from being innovative in some agencies because the Inspector General would denounce them for non-compliance (Trodden, 1996).

In addition, demographic variables including race, gender, education, tenure, age, and income level are important factors in previous individual and organizational performance management studies. Many studies have found that these variables may influence performance in studies of job satisfaction (Bender & Heywood, 2006), student performance (Booker, Gilpatric, & Gronberg, 2007; Pitts, 2005; Sabharwal, 2014), organizational commitment (Cho & Mor Barak, 2008), and career outcomes (Greenhaus, Parasuraman, & Mormley, 1990). Race and gender are the crucial variables on decision making, interpersonal relations, workforce cohesion, and even performance (McKay, Avery & Morris, 2009). For example, one study has found that white and male managers felt more accepted in their organizations, and perceived themselves as receiving higher ratings from their supervisors on the job performance (Greenhaus, Parasuraman, & Mormley, 1990). In addition, diverse education programs or levels had positive effects on performance (Bender & Heywood, 2006; Ely, 2004) while tenure and age were negatively related to overall job satisfaction (Ely, 2004; Lee & Wilbur, 1985).

Challenges of Performance Management and Measurement in the Public Sector

To have meaningful impact, federal agencies have adopted various management skills and measurement indicators to help them determine whether their missions are accomplished, and their goals achieved (Behn, 2003; Radin, 2006; Rainey, 2014; Starling, 2010). Despite strong theoretical and empirical evidence supporting the efficacy of performance management and measurement, many scholars have noted challenges that must be overcome (Moynihan, 2008; Moynihan & Pandey, 2004; Radin, 2006; Talbot, 2010).

Since most performance measures are designed and implemented by managers in a top-down fashion (Brown & Potoski, 2003), a growing number of studies confirm that management, especially the manager's role, matters (e.g., Andrews, Boyne, Meier, O'Toole, & Walker, 2005; Brewer, 2005; Chun & Rainey, 2005; Lan & Rainey, 1992). Therefore, a manager's inappropriate decision from lack of knowledge or information makes a critical difference for the success of performance measurement and performance management (Julnes, 2008; Hatry, 2006). Managers' professional knowledge generates a better understanding of reliable indicators for measurement (Banks & Murphy, 1985). Moreover, managers engage in strategic decision-making by setting goals and objectives and this can result in low measurement coherence if managers stray from agency goals and objectives. The selection of performance measures is dynamic and organizational success depends on the managers' decisions about what should be measured and how, when, and by whom it should be measured. Managers are not cognizant of all performance information and their responsibilities sometimes change in ways that affect their interests and concerns. The resulting low congruence leads to poor measurement validity.

Furthermore, due to the lack of objective data on organizational performance and performance measurement, many scholars have used perceptual performance measures (Fernandez & Moldogaziev, 2011; Jung & Lee, 2013; Moynihan & Pandey, 2004). Most

importantly, scholars point out performance measurement challenges in Public Administration since many studies have focused solely on perception data. Although perfect measures hardly exist, researchers must be aware that perceptions may include bias inflating the results (Favero & Bullock, 2014; Meier and O'Toole, 2013b). Common source bias exists based on certain patterns or tendencies that "respondents apply, or that impact their responses, systematically across different measures when answering a survey," creating a high correlation between variables (Jakobsen & Jensen, 2015, p. 5). Finding an independent source of data or using two distinct sets of survey respondents may help ensure that the results are not affected by common source bias (e.g., Meier & O'Toole, 2013b). Therefore, Favero and Bullock (2014) suggest that scholars design their research by finding an independent source of data in order to reduce or mitigate the possibility of obtaining biased results when using perceptual measures. In this vein, based on the challenges of performance measurement, comparing subjective and objective archival data is meaningful (Starbuck & Mezias, 1996), and some empirical studies have tried to adopt multiple measurement strategies (Cameron & Whetten 1983; Rainey, 2014).

With regard to performance measurement, most outcomes/impacts are difficult to measure. Setting and measuring outcomes are important because they make agencies move towards their ultimate objectives by managing performance. The ultimate goals of public services that are multi-dimensional, social, and collective in nature, such as development of community sustainability, are difficult to measure compared with output and process measurement, such as hours of operation, which is relatively easy to measure (Bouckaert & Balk, 1991; Julnes, 2008; Hatry, 2006; Poister, 2008). Meier and O'Toole (2006) add that Public Administration is a science supporting the normative view and should be responsive to political authority and democratic-constitutional principles, such as transparency, equity, and

accountability (Meier & O'Toole, 2006). In the similar vein, another major reason for the challenges is that agency goals are not always clear. Many studies have found that goal ambiguity linked to unclear criteria has a negative influence on organizational performance in public agencies (Chun & Rainey, 2005; Jung, 2013). Since agency goals contain competing values and target goals (Rainey, 2014), explicit organizational goals help agencies achieve their mission by identifying what they are doing and what they should be doing (Chun & Rainey, 2005; Weiss & Piderit, 1999).

Performance Management as Creating Public Value

One argument for performance measurement and management is rooted in 'creating public value' (Moore, 1995). Public services are not only considered as products, but also as adding value, such as equity, equality, and building social capital. Jørgensen and Bozeman (2007) consider values as ideals for producing public services in the public sector because those public values provide direction to public employees for achieving the desired results of policy (Andersen, Brewer, & Leisink, in press). According to Horner, Lekhi, and Blaug (2006), public officials demonstrate that one of the key principles of measuring public sector performance is that public value measurement blends with strategic goals of the organization (e.g., de Buijn & Dicke, 2006; Poister & Streib, 1999). Even though value creation in the public sector involving public goods and services is difficult, public managers have been encouraged to measure various dimensions of their outputs such as efficiency, service quality, customer satisfaction, and service equity. This starts with identifying important values in public services in order to demonstrate that public resources have been used well (Andrews & Boyne, 2010; Kelly, 2005; Smith & Huntsman, 1997).

Previous research has tended to emphasize the principles of administrative efficiency not only in the Public Administration field but also in performance management studies, while neglecting measures that are more difficult to quantify, such as social equity, fairness, and justice (Andrews & Van de Walle, 2013; Hatry, 2002; Ingraham, 2005; Piotrowski & Rosenbloom, 2002; Pollitt, 2002). This limited approach to measuring performance is not enough to guide public administration practice, and it can lead to incomplete conclusions about organizational effectiveness (Kaplan & Norton, 1992; Yang & Holzer, 2006). Indeed, program administrators and managers are confronted with the need to create a multidimensional construct of performance to tap the multiple values and goals that are embedded in public programs (Boyne & Dahya, 2002; Brewer & Selden, 2000; Brewer & Walker, 2010; Chun & Rainey, 2005; Pandey & Rainey, 2006; Rosenbloom, 1983).

Recently, the field of Public Administration has responded to this need for a fuller spectrum of performance measures by emphasizing public values (Hill & Lynn, 2009; Lynn Jr., 2006; Moore, 1995), although the definition of public value is ambiguous and the approaches for understanding what it means vary (e.g., Jørgensen & Bozeman, 2007; O'Flynn, 2007; Stoker, 2006). Buchanan (1975) contends that conflicting values inherently exist in a public organization's goals (Downs, 1967; Rainey, 2014). He argues that not only is it extremely difficult to measure each value because of its unclear definitions and intangible goals, but most programs are comprised of a complex combination of goals related to efficiency, effectiveness, and equity, which often conflict (Brown, Potoski, & Van Slyke, 2006; Hoggett, 2006; Hood, 1991; Jørgensen & Bozeman, 2007; Radin, 2006; Rainey, 1989).

The basic values of public service have been discussed by numerous scholars. These values include responsibility (Finer, 1941; Friedrich, 1940), responsiveness (Cohen, 2008;

Lipsky, 1980; Yang & Callahan, 2007), accountability (Dubnick, 2005; Dubnick & Frederickson, 2009; Koliba, Mills, & Zia, 2011; Koppell, 2008; Moe & Gilmour, 1995), commitment (Stoker, 2006), fairness/equity (Brewer & Selden, 2000; Frederickson, 1990, 1997; Selden, Brewer & Brudney, 1999; Stoker, 2006), transparency (Ljungholm, 2015; Yang & Holzer, 2006; Yang & Rho, 2007), ethics and morality (Frederickson, 1999), diversity (Kellough & Naff, 2004; Pitts, Hicklin, Hawes & Melton, 2010), citizenship and the public interest (Brown et al., 2006; Denhardt & Denhardt, 2015), justice (Greenberg, 1990; Linna, Elovainio, Van den Bos, Kivimäki, Pentti, & Vahtera, 2012), representation (Sowa & Selden, 2003; Stoker, 2006), discretion (Moynihan, 2008; Sandfort, 2000), expertise and professionalization (Noordegraaf, 2007), political neutrality (Rourke, 1992), efficiency (Andrews & Van de Walle, 2013; Greiling, 2006), effectiveness (Rainey & Steinbauer, 1999), productivity (Bouckaert, 1992; Taylor, 1911), professionalism (Mosher & Stillman, 1982), and dignity/respect (Brewer, 2013).[2] Obviously some values may be more important than others, particularly in a given situation. Several prominent values are discussed in more detail below.

Public sector performance measurement has focused on traditional values that are relatively easy to gauge such as efficiency (Downs & Larkey, 1986; Greiling, 2006; Kirlin, 2001). Increasing the efficiency of public services has been high on the agenda of the promoters of New Public Management (Andrews & Van de Walle, 2013). Hatry (2006) defines efficiency measures as "the relationship between the amount of input and the amount of output or outcome of an activity or program" (p. 13). Accordingly, many studies focus on economic efficiency to evaluate public policies and programs (Downs & Larkey, 1986; Langbein, 1980). Efficiency is also frequently explained by the behavioral predictions of the utility-maximization model with

[2] This list is illustrative but not exhaustive.

consideration of multidimensional perspectives, such as production, costs, and results (Friedmann, 2002; Weimer & Vining, 2015). Yet Hatry (2006) warns that the most cost-effective solution might not be the right plan for a particular program.

Some scholars point out that too little attention has been given to service quality (Berman, 1997; Bok, 2001), satisfaction (Kelly & Swindell, 2002; Kelly, 2005), and social equity (Frederickson, 1990, 1997; Merget & Renee, 1982; Rich, 1979; Rosenbloom, 1983) due to unclear definitions that make the concepts difficult to measure (Boyne, 2003; Frederickson, 1996; Jennings, 2005). Many scholars have acknowledged the importance of service quality and urge more attention to managing service quality in order to serve citizens better (Denhardt & Denhardt, 2000). Service quality measures may be based on aspects such as thoroughness, accuracy, convenience, and well-mannered and professional treatment of clients (Hatry, 2006). Besides, customer service satisfaction measures as performance outcomes have been increasingly important (Vermeeren, Kuipers, & Steijin, 2011). As services are inherently intangible, customer perceptions of the expertise of the service are directly linked to loyalty (Brady & Cronin, 2001) and political legitimacy (Dahl 1971). In addition, social equity has a significant component in public policy. There are several types of equity: adequacy, end-results equity, horizontal equity, vertical equity, process equity, distributional equity, and equal access (Langbein, 1980; Stone, 2002; Svara & Brunet, 2004; Tresch, 2008). Even though there is no agreement on what an equitable distribution is, equity cannot be ignored because one of the key roles of government is to distribute public goods and services (Lasswell, 1936; Stone, 2002; Svara & Brunet, 2004).

The number of performance measurement indicators that reflect a particular public value will likely increase over time for two reasons. First, value content often includes complex

democratic principles that cannot be measured by a single concept or small number of indicators. For example, the concept of social equity includes horizontal equity (e.g. providing equal resources to individuals based on the grounds such as race and gender) and vertical equity (e.g. providing equal resources to individuals based on social class or income). Thus, public values require diverse indicators that include various perspectives in order to capture their complexity (Pidd, 2012). Second, over time, performance measurement becomes more precise about what agencies want to assess and more explicit about what metrics are used for assessing it. Many studies have suggested that multiple measures are more useful in capturing organizational or program goals (Pidd, 2012) because diverse metrics improve measurement accuracy and precision through triangulation and repeated measures. Therefore, governments often develop deliberate, explicit, and numerous indicators to capture diverse value-added components (Challis, Clarkson, & Warburton, 2006; Pidd, 2012). This could explore research hypotheses H1.

Moreover, these values are found at each stage of the policy process model (input-process-output-outcome) model (Brown, 1996; Fountain & Roob, 1994; Hatry, 2006). Quinn and Rohrbaugh (1983) also point out that an organization may emphasize different values in various organizational settings and stages in the organization's life cycle. As each stage of the policy process model has different characteristics as discussed above, different public values are frequently expressed in certain stages. Value content is present in every stage of the policy process model. Performance indicators in the process stage are usually expressed in quantitative terms for detecting progress or lack of progress to achieve higher-level results (Ömürgönülşen, 2002). Bird et al. (2003) contend that performance indicators for the process stage should be straightforward and unambiguous so as to allow managers to easily measure process. The output stage of the policy process model is related to the amount of services or products delivered and it

is more easily measurable for quantifying performance. According to Kelly, Muers, & Mulgan (2002), process and output indicators are more specific than outcomes, which are often broader and vaguer. Since outcome represents the most ambitious results with the highest-level objectives, effective outcome indicators usually include abstract goals with qualitative measures. Outcomes involve more comprehensive components that unfold over time. Therefore, value content is mainly expressed in the outcome stage, which is directly related to the impact of public services.

Accordingly, this study will emphasize the value content of the performance measurement criteria of efficiency, customer service satisfaction, service quality, and social equity that appear in federal employee surveys and annual performance and accountability reports of federal agencies, and it will sort those values into the various stages of the policy process model to enable a finer-grained assessment. This latter step will provide important insights on agency efforts to integrate the various performance values into their performance management process overall. This leads to formulate research hypotheses H2. One expectation is that patterns in 2013 and 2017 are somewhat uneven and inconsistent, given the problematic nature of performance measurement as characterized in the literature and based upon anecdotal accounts. On the other hand, value content may be relatively consistent across 2013 and 2017 because it evolves incrementally and may be minimized by institutional stability and the agency's culture. This study adopts the former expectation. The rationale comes from the experience of different presidential administrations between 2013 and 2017.

According to official documentation (e.g., Trump's President's Management Agenda), there has been a continuation of attention to performance measurement as a tool for performance management, which started in the Clinton administration. One of the major themes in the Trump

administration is, however, that each program should identify more detailed performance metrics by employing performance information. This difference improves the agency's efforts to achieve precise measurement which should reduce the gap between objective measures and managers' perceptions. Moreover, according to the Trump administration's agenda, the key driver of reform in performance management is the modernization and development of information technology, including data security and an integrated data strategy (e.g., collecting common indicators on cross-agency goals and priorities). Agenda of Modernizing Government for the 21st century has been attempted to improve data accuracy, detailed performance information which can reduce the gap between objective measures and managers' perception. Accordingly, this study examines whether there are differences between objective measures and managers' perception of value content for assessing public service performance. Thus, this can be developed research hypotheses H3. Thus, the following hypotheses will be tested:

H1. Value content will be more frequently expressed in the 2017 annual agency report than in the 2013 report.

H1a. The efficiency value will be more frequently expressed in the 2017 annual agency report than in the 2013 report.

H1b. The service quality value will be more frequently expressed in the 2017 annual agency report than in the 2013 report.

H1c. The customer service satisfaction value will be more frequently expressed in the 2017 annual agency report than in the 2013 report.

H1d. The social equity value will be more frequently expressed in the 2017 annual agency report than in the 2013 report.

H2. Value content will be more frequently expressed in the outcome stage than in the input and output stages in the annual agency report.

H2a. The efficiency value will be more frequently expressed in the outcome stage than in the input and output stages in the annual agency report.

H2b. The service quality value will be more frequently expressed in the outcome stage than in the input and output stages in the annual agency report.

H2c. The customer service satisfaction value will be more frequently expressed in the outcome stage than in the input and output stages in the annual agency report.

H2d. The social equity value will be more frequently expressed in the outcome stage than in the input and output stages in the annual agency report.

H3. The gap between archival-based and perceptual measures of value content will be less in 2017 than in 2013.

H3a. The gap between archival-based and perceptual measures of efficiency will be less in 2017 than in 2013.

H3b. The gap between archival-based and perceptual measures of service quality will be less in 2017 than in 2013.

H3c. The gap between archival-based and perceptual measures of customer service satisfaction will be less in 2017 than in 2013.

Some additional hypotheses are formulated based upon how proximity can affect performance management and performance measurement processes in the public sector. This topic is discussed in the next chapter.

CHAPTER 3

PROXIMITY THEORY

As mentioned in Chapter 1, performance management involves managerial perceptions of organizational aspirations, and performance management systems require managers to invest the time and effort needed to identify the appropriate performance measures. They also require cognitive efforts from decision-makers to set the targets, to establish performance metrics, to evaluate outputs and outcomes, and to adjust their goals based on evidence from the feedback. Cognitive efforts are the process by which people recognize information, feelings, and action based on experiences. Therefore, "the quality or accuracy of a person's perceptions has a major impact on his or her responses to a given situation. Perceptual responses are also likely to vary between managers and subordinates" (Schermerhorn, Osborn, & Hunt, 2007, p. 32). Actors perceive a situation differently because of their different positions and diverse individual needs (Schermerhorn et al., 2007). This is the dynamic described in Miles law: "where you stand depends on where you sit."

Surprisingly, little research in Public Administration has been devoted to examining the coherence between managers' perceptions of agency performance criteria and objective measures of those criteria at the agency level. This is a form of performance measurement validity and it is important to examine the degree of coherence and its causes and effects. Hence, inspired by proximity theory, this study will demonstrate how individuals (managers and

employees) construe events as a function of distance, and how this can influence their judgments, decision-making, and subsequent behavior, especially regarding performance measurement.

Definition of Proximity Effects

The concept of proximity comes from the Latin *proximus* or "nearest" (Kreger, 2008). Proximity is defined as "being close to something" (Knoben & Oerlemans, 2006, p. 71). The concept of proximity is widely discussed in a proxemic theory introduced by anthropologist Hall (1966). From the beginning of the nineteenth century, the proximity concept has been studied in many social science disciplines (Knoben & Oerlemans, 2006). Proximity is, for example, an important variable in industry science and innovation (Boschma, 2005; Gilly & Torre, 2000; Oerlemans, Meeus, & Boekema, 2001), regional economic development (Gilly & Torre, 2000; MacKinnon, Cumbers, & Chapman, 2002; Wilhelm & Ritz, 2003), inter-organization collaboration (Sternberg, 1999), organizational science (Boschma, 2005; Knoben & Oerlemans, 2006), psychology (IJzerman & Semin, 2009; Monge & Kirste, 1980; Stern, Cole, Gollwitzer, Oettingen, & Balcetis, 2013), and more recently, economics (Torre & Zuindeau, 2009; Zimmermann, 2008), health and physical activity (Mowen, Orsega-Smith, Payne, Ainsworth & Godbey, 2007), sociology (Rivera, Soderstrom, & Uzzi, 2010; Urry, 2002), and marketing and business (Bergadaà & Del Bucchia, 2009; Wynder, 2014).

Understanding the proximity concept and its dimensions is important because different forms of proximity can affect managers' perceptions on performance measurement and management systems. While many scholars have been paying more attention to the concept of proximity across the social sciences, it has received little or no attention in Public Administration, especially in performance measurement and performance management. The

concept of proximity has not been clearly understood while its importance seems to be evident for organizations. The proximity concept is synonymous with presence, co-presence, cohesiveness, and propinquity. Presence, which is mostly studied in the computer science field, is defined as the "perceptual illusion of non-mediation" (Lombard, Reich, Grabe, Bracken, & Ditton, 2000, p. 77) or "a psychological state in which the virtuality of experience is unnoticed" (Lee, 2004, p.32). Co-presence refers to the "sense of being together with others in a mediated — either remote or virtual — environment" (Zhao, 2003, p.445); while cohesiveness, which is a group level concept, refers to "the overall attraction or bond amongst members of a group" (Wellen & Neale, 2006, p. 168). There are two kinds of cohesiveness: social and task cohesiveness. Social cohesiveness describes group attractions based on social relations while task cohesiveness is based on a shared commitment for achieving group goals. Similarly, the propinquity effect has been defined as the "higher likelihood of the formation, maintenance, and strength of social network ties in geographic proximity" (Mevners, Barrot, Becker & Goldenberg, 2017, p. 52) and it is widely used in political science, social psychology, and sociology (Burgoon, Bonito, Ramirez, Dunbar, Kam, & Fischer, 2002; Newcomb, 1956). The propinquity effect is mainly used for explaining interaction and communication based on geographical proximity. This concept is, however, limited and allied to geographical proximity.

These concepts seem related to proximity, but they are conceptually distinct. While they help predict how relationships between closeness and outcomes develop in organizations, there are two reasons for using the concept of proximity instead. First, early theories of presence and co-presence may not account for socio-cultural dimensions and current patterns of behavior. Concepts such as presence and co-presence have faced challenges to explore detailed types of closeness while the concept of proximity includes various dimensions of closeness. For example,

Jones (1991) breaks down proximity into cultural, social, psychological, and physical dimensions, while Boschma (2005) provides different forms of proximity— cognitive, organizational, social, institutional, and geographical proximity. Second, propinquity represents mainly physical closeness. Since proximity is found at different levels of analysis, different labels are used to denote various dimensions of proximity that can be overlapping or interconnected, thus strengthening each other (Boschma, 2005; Knoben & Oerlemans, 2006). For example, geographical proximity can refer to groups of firms in a geographical unit (relational definition, agglomerations) or as the geographical distance between two collaborating firms (spatial definition).

Even though researchers have studied proximity in many different fields, some scholars argue that distance is no longer relevant in a fast-changing environment with emerging new technologies (Olson, Teasley, Covi, & Olson, 2002). Others suggest that more sophisticated research is needed to explore proximity, emphasizing the distinction between proximity dimensions (Lackey & Kaczynski, 2009; Scott, 1999). Thus, this study uses proximity as a multi-dimensional concept to investigate the impact of managers' closeness to performance-related activities on the performance measurement process.

There are two broad kinds of proximity: objective proximity and subjective proximity (Cole, Balcetis, & Dunning, 2013; Gibbs, Costa Lima, & Francozo, 2004; Lackey & Kaczynski, 2009). Both forms of proximity can stimulate goal-relevant behavioral responses, encouraging motivation and action (Cole et al., 2013). Objective proximity is objectively measured or calculated as spatial distance while subjective proximity is defined as actors' perceptions of closeness or nearness to persons, organizations, or systems (O'Leary, Wilson, & Metiu, 2014; Wilson, O'Leary, Metiu, & Jett, 2008). Objective proximity comprises geographical (Cramton,

2001; Jones, 1991; Maskell, 2001; O'Leary et al., 2014; Torre & Rallet, 2005), and subjective proximity includes psychological and cognitive (Cohen & Levinthal, 1990; Wuyts, Colombo, Dutta, & Nooteboom, 2005), organizational (Meisters & Werker, 2004), institutional (Kirat & Lung, 1999), social (Bradshaw, 2011; Jones, 1991), goal (Bandura & Schunk, 1981; Jhang & Lynch, 2014; Kivetz, Urminsky, & Zheng, 2006), cultural (Gill & Butler, 2003; Huang, Chen & Lin, 2013; Jones, 1991), and technological proximity (Greunz, 2003). Subjective proximity infers cognitive or emotional closeness– as implied when a person says: "I know his/her mind." Previous studies have concluded that a decision maker who is psychologically close to colleagues or his/her workplace is likely to show greater responsibility for relationships and outcomes, thus facilitating effective communication. Subjective proximity means sharing the same knowledge base and expertise, increasing success in communicating and understanding new information (Boschma, 2005; Boschma & Lambooy, 1999). It also There are many dimensions of subjective proximity and this study explores the characteristics of geographical, organizational, institutional, social, cultural, psychological, informational, and goal proximity.

Geographical proximity refers to the spatial closeness or physical distance between a decision-maker and those the decision will affect (Wynder, 2014). Many scholars equate objective proximity with geographical proximity, which refers to the spatial or physical distance between actors (Boschma, 2005). Geographical proximity can be illustrated in many ways—the number of hallways and buildings separating work group members, different locations where people work, and work sites away from headquarters (e.g., Finholt, Sproull, & Kiesler, 2002). Maskell (2001) states that co-location in transparent clusters helps to maintain stable and informal relations among actors because geographical proximity facilitates interaction and cooperation that leads to more personal and embedded relationships based upon trust (Dahl &

Pedersen, 2004; Lam, 1997; Malecki & Oinas, 1999). Cramton (2001) also examines collaboration with a physical dispersion of group members, and she attributes failures in building mutual knowledge to the lack of face-to-face communication that may lead to unevenly distributed information and differences in the salience of information. Reed and Phillips (2005) suggest that geographical proximity to an exercise facility may reduce psychological barriers to exercise. In their research, geographical proximity to exercise facilities is identified as the distance between the facilities that students use and their residences. The authors found that geographical proximity to exercise facilities based on intensity, frequency, and duration of physical activity positively affects individual decisions to exercise. Most studies in the past have defined objective proximity as geographical proximity. Recently, however, exploring only objective proximity and distance in physical terms has been regarded as an incomplete view of how people experience proximity (Boschma, 2005; Knoben & Oerlemans, 2006; Wilson et al., 2008). Therefore, recent analysis considers geographical proximity not only as spatial distance which is objectively calculated (e.g., number of miles) but also as individual perception of physical nearness (Mencl & May, 2009; Torre & Rallet, 2005). For example, Hess and Almeida (2007) assessed the relationship between proximity to transit stations and residential property values in New York. They explored objective proximity (actual walking distance to stations) along with perceived proximity to stations and found that subjective proximity effects that increase accessibility benefits are positive in high-income station areas and negative in low-income station areas. According to O'Leary et al. (2014), geographical proximity has little effect on the quality of team relationships, but subjective proximity has significant effects on that variable because individuals can perceive higher levels of proximity when they have frequent meetings and intense communication, even if they are not physically close to each other. In this

sense, some scholars have argued that subjective proximity is much more than just "being there" (Hollan & Stornetta, 1992, p. 120) while some others have declared that physical distance is dead (Cairncross, 2002).

Organizational proximity includes intra-organizational and inter-organizational relations of members. Regarding intra-organizational relations, organizational proximity refers to the level of interdependencies between organizations "connected by a relationship of either economic or financial dependence/interdependence between member companies of an industrial or financial group, or within a network" (Kirat & Lung, 1999, p. 30). In terms of inter-organizational relations, organizational proximity refers to the same space of relations, which includes similarity between individuals in the organization by sharing the same information and space (Gilly & Torre, 2000). In this perspective, organizational proximity refers to "the proximity between employees of a multi-plant firm who identify with each other as a result of belonging to the same firm and sharing knowledge of the firm's specific routines" (Schamp, Rentmeister, & Lo, 2004, p.609). Organizational proximity focuses more on the structural equivalence of actors (e.g., Rice & Aydin, 1991) and characteristics of the network (e.g., Oerlemans & Meeus, 2005). Some studies focus on organizational proximity as the characteristics of the network rather than on the characteristics of relationships between two agents, while others pay attention to specific relationships – the so-called dyadic level (e.g., Wilkof, Brown, & Selsky, 1995). Wilkof et al. (1995) focused on characteristics of the organization and its network structure and found that firm B with an informal, decentralized, and non-bureaucratic structure used a win-win perspective mechanism to achieve success, compared to firm A with a hierarchical and bureaucratic structure. Overall, most studies have found that organizational proximity improves

outcomes, such as organizational performance, when individuals are exposed to similar organizational contexts.

Institutional proximity refers to "sets of common habits, routines, established practices, rules, or laws that regulate the relations and interactions between individuals and groups" (Edquist & Johnson, 1996, p. 46). Institutional proximity is high when actors share norms, rules, and laws (Davids & Frenken, 2018). Institutional proximity includes two forms: co-location with the same territory (Boschma, 2005) and joint participation under the same social sub-system, such as industry or government (Ponds, Van Oort, & Frenken, 2007). Moreover, institutional proximity is studied on two different levels of analysis, since institutions consist of both informal constraints (i.e., sanctions, taboos, customs, traditions, and codes of conduct) and formal rules (i.e., constitutions, laws, legislative conditions, and property rights) (Johnson, 2001; Zukin & DiMaggio, 1990). In other words, institutional proximity denotes the level of congruence between organizations; thus, it is found at the macro-level (Freel, 2003). Institutional proximity facilitates collective learning by sharing knowledge transfer among agents based on models, norms, and procedures (Capello, 1999; Kirat & Lung, 1999). Therefore, institutional proximity is conflated due to the level of similarity of the norms and routines between organizations (Knoben & Oerlemans, 2006).

Social proximity refers to socially embedded closeness relations between partners in interaction based on trust (Boschma, 2005; Matthews & Matlock, 2011). Since social proximity is affected by individual or group feelings based on their interaction, social proximity has a strong positive correlation with trust (Boschma, 2005; O'Leary, Wilson, & Metiu, 2014). It requires trust-based social relationships to encourage an attitude of communicative rationality that stimulates interactive commitments based on the friendship and kinship that are built based

on emotional bonds (Boschma, 2005; Lundvall, 1993; Uzzi, 1997). Simply put, social proximity is personal proximity (Schamp et al., 2004) or relationship proximity (Coenen, Moodysson, & Asheim, 2004). Amin and Wilkinson (1999) have found that proximity between actors has a significant influence on stimulating interaction and performance. Sommer (1962) also found that closer interpersonal distances are formed by informal conversational settings and face-to-face seating arrangements.

Cultural proximity is conceptualized as people's perception of similarities and differences between cultures (Knoben & Oerlemans, 2006; Wang, 2016). Based on the meaning of culture, which is a pattern of thought, feelings, and behaviors, culture is shared and accepted by a group, binding members together and providing similar interpretations of situations (Knoben & Oerlemans, 2006; Pettigrew, 1979; Wilkof, Brown, & Selsky, 1995). A study conducted by Babiker, Cox, and Miller (1980) developed an index that measured the perceived cultural similarities based on such characteristics as "climate, food, language, clothes, religion, educational level, material comfort, family structure and family life, courtship and marriage, leisure activities, and intergroup conflict" (p.109). Individuals who have similar features, such as race, ethnicity, age, religion, education, and occupation, are more likely to interact and connect to those they perceive as similar to themselves (Smith-Lovin, 2003). Cultural differences between partners have a negative impact on the level of the trust relationship, increasing conflicts (Bruneel, Spithoven, & Maesen, 2007). According to Kirkman, Lowe, and Gibson (2006), the most popular application of cultural proximity is a cultural index which includes five dimensions: individualism, power distance, uncertainty avoidance, masculinity, and Confucian dynamism (Hofstede, 1980; Hofstede & Bond, 1988). These dimensions include the degree of closeness within/between social frameworks, the level of inequality of power distribution among

agents, and the use of adequate measures to prevent ambiguous relations and situations to secure principles and values in society (Kirkman et al., 2006).

Psychological proximity is a relational attribute developed by Nooteboom (1990, 2000). Based on one's direct experience, the actor's perception, interpretation, and understanding are determined (Wuyts et al., 2005). Accordingly, it can be used to describe an individual's experience of events or other people because the reference for psychological distance is oneself, and distance is captured by perceived proximity to self (Trop & Liberman, 2010). Therefore, psychological proximity is an important factor for risk perceptions and experiencing affect (Lent Sungur, Kunneman, & Das, 2017). Previous research has shown that an individual's perception of risk for negative events decreases with increased psychological proximity (e.g., Williams, Stein, & Galguera, 2014; Davis, Gross, & Ochsner, 2011). For example, Rosenbaum, Schuck, Graziano, and Stephens (2007) scrutinize the relationship between police programs and community performance; the public scores police programs highly when they perceive a good relationship between police and their neighborhoods, feeling less fear of crime and social and physical disorder.

Informational proximity, which is based on shared professional experiences and knowledge, facilitates the process of the relationship that leads to new products or services (Knoben & Oerlemans, 2006; Tushman & Anderson, 1986). It accelerates the acquisition and development of trained knowledge and technologies between actors or between actors and knowledge (Knoben & Oerlemans, 2006; Tremblay, Klein, Fontan, & Rousseau, 2003; Zeller, 2004). Compared to psychological proximity, which focuses on the issue of 'how' actors interact, informational proximity pays attention to 'what' they exchange (Boschma, 2005). It is related to the concept of 'absorptive capacity' (Boschma, 2005; Lane & Lubatkin, 1998), which

refers to the "ability to recognize the value of new, external knowledge, assimilate it and apply it to commercial ends" (Cohen & Levinthal, 1990, p. 128). Organizations with high informational proximity tend to acquire high levels of expert knowledge, and they produce more innovation than organizations with low informational proximity. On the other hand, lack of informational proximity between organizations leads to a low level of collaboration.

Goal proximity has been examined mostly in the fields of business management, behavioral study, and psychology. Behavioral researchers observed that rats in a maze ran faster as they approached their goal (Hull, 1932). This phenomenon has also been studied in humans, assessing human tendencies concerning goal achievement (Heilizer, 1977). Many studies agree that goal proximity is related to performance. For example, Bandura and Schunk (1981) set up seven sessions for children with low subtraction skills. The first group had a proximal goal of completing one set in each session while the second group had a distant goal of completing all sets by the end of the last session. Based on the results, the authors assert that short-term goals are more effective due to more frequent assessment, stimulating improvement of self-confidence when the children achieve goals. Since short-term goals are more flexible and controllable, they can be raised or lowered to maintain the optimal challenge level. On the other hand, Kirschenbaum (1985) supposes that long-term goals facilitate performance improvement. Long-term goals perceived as being a long distance away provide direction and motivational effects for achievement striving. Moreover, regardless of goal type, many studies have found that goal proximity is a crucial factor for increasing efforts towards achieving a goal (Kirschenbaum, 1985; Kivetz, Urminsky, & Zheng, 2006). Martin and Tesser (1996) examined whether the relationship between proximity of goal setting and the type of goal setting interacts with goal direction on performance. They found a significant mediation effect of goal proximity for better

performance (e.g., Jhang & Lynch, 2014; Tenenbaum, Pinchas, Elbaz, Bar-Eli & Weinberg, 1991). Balcetis, Cole, and Bisi (2015) also provided evidence that a high level of goal proximity increases goal achievement, as being close to a goal increases goal-relevant action.

Individual-Level Effects of Proximity and Performance

Many studies of proximity have addressed performance-related topics in other social science disciplines, but neither the concept nor the relationship have been studied systematically in Public Administration. Some studies have found a positive relationship between proximity and outcome because when individuals are close to the desired end state, they believe they are capable of achieving social, physical, and cognitive tasks (Cheema & Bagchi, 2011; Kraut, Fussell, Brennan, & Siegel, 2002). Buchanan (1975) contends that managers who feel they make real contributions to organizational success are more likely to develop commitment than those who do not contribute. Managers and leaders who are involved in a program have a vested interest in employee success, organizational achievements, and managing performance. Cha, Park, and Lee (2014) explored the association between subjective proximity and teamwork quality. They found that the perception of closeness positively mediates relationships between team members and team outcomes based on collaborative work environments because team members feel psychologically connected to each other. Cole et al. (2013) also assert that both objective and subjective proximity improve motivation and performance. When an individual's perception of the distance between actors and object has decreased, they are more likely to be encouraged to act, regardless of the level of spatial separation. Balcetis et al. (2015) analyzed whether proximity affects the anticipated outcome, namely appraisals of desirability (e.g., Ajzen,

1985; Atkinson, 1957; Hull, 1943), and found that subjective proximity stimulates appraisal feasibility but does not affect appraisal desirability.

Other studies of proximity have reported inconsistent and conflicting findings regarding the effects of subjective proximity and performance (Martins, Gilson, & Maynard, 2004; Wilson et al., 2008). For example, some studies have found a negative relationship between physical distance and desired outcomes in work (e.g., Allen, 1977; Festinger, 1951; Short, Williams, Christie, 1976). Other studies have even concluded that perception does not predict performance (Witt & Dorsch, 2009).

In sum, there are several limitations of prior studies. First, few studies have focused on the influence of objective and subjective perception on a group's or individuals' performance (Witt, 2011). Second, while the proximity concept seems to be an important, emerging concept in diverse fields of social science, it has received only scant attention in the field of Public Administration with limited descriptions based upon the broad concept of proximity. Therefore, this study aims to help fill this gap in the literature by applying the concept of proximity in the Public Administration field. More specifically, this study will determine the impact of proximity on performance measurement validity in U.S. federal government agencies.

Proximity and Performance Measurement Validity

As covered in Chapter 1 and 2, managers establish organizational goals and objectives based on past aspirations, performances, inherited attributes, inertia, organizational capabilities, and managerial perceptions of need (Ansoff, 1987; Shinkle, 2012). Since an agency's annual report reflects the managers' perceptions of needs, it is important to examine whether the agency's annual report and the managers' perceptions of the organizational goals are coherent.

Thus, performance measurement validity means measurement coherence or congruence between the measures in the agency's annual report and the managers' perception of measures in the agency's annual report.

Recent scholars have examined employee perceptions of performance and effectiveness in the public sector (Fernandez & Moldogaziev, 2011; Jung & Lee, 2013; Moynihan & Pandey, 2004; Stern et al., 2013; Witt, 2011). Managers' cognitive feelings and perceptions may be predisposed by diverse forms of proximity resulting from their positions, roles, and interactions (Witt, Linkenauger, & Proffitt, 2012). Two broad forms of proximity may be associated with performance measurement validity: objective and subjective. According to the literature review, objective proximity includes actual geographical distance and actual involvement in decision making that is objectively calculated or measured. Managers' actual involvement in decision making and administrative engagement in performance management activity can reduce potential biases (Liu, Yang, & Yu, 2015). When managers are involved in decision making and are actively engaged and participating, they are more likely to possess and report accurate information (Huber & Power, 1985).

Since many scholars have employed subjective measures, DeNisi and Pitchard (2006) clarify that measurement validation issues, including the rater's perception, motivation, and ability to measure accurately, are important for ensuring fair measurement of performance (Murphy & Cleveland, 1995). Meier and O'Toole (2012) propose that managers' subjectivity distorts their perceptions of performance, addressing empirical problems with administrators' subjective assessments of performance. The authors found that the perceptual biases in self-assessments of performance are correlated to measurement error, which can be a problem no matter how the independent variables are measured, and can lead to spurious correlations. The

perception of individuals' position and work conditions may affect performance. Brewer (2011) contends that individuals' position may relate to not only cognitive biases but also their perceptions of performance. For example, managers with a high level of familiarity with the process of setting and prioritizing an agency's goals, objectives, and performance measures tend to have more favorable views of their performance system, which can cause a decrease in accuracy of outcome. Non-managers tend to focus on their day-to-day tasks rather than their agency's goals. (Banks & Murphy, 1985; DeNisi, Cafferty, & Meglino, 1984; Klimoski & Inks, 1990; Mero, Guidice, & Anna, 2006).

Based on the findings of previous studies, some studies have found that objective proximity increases accuracy while subjective proximity may be associated with cognitive biases. On the other hand, both objective and subjective proximity increase accuracy. Therefore, reliable measures cannot be achieved by using only objective or subjective measures, but could be achieved by comparing subjective performance measures with objective performance measures to attain a higher level of measurement validity (Wang & Gianakis, 1999). The distinction between subjective and objective performance measures lies in "whether the measure is based on empirical observation or, alternatively, on beliefs, perceptions or attitudes" (Wang & Gianakis, 1999, p. 538). For example, a local utility department provides safe and reliable drinking water for the community. A number of emergency service alerts are empirically counted and observed, which are objective measures. On the other hand, the percentage of residents who feel safe and the percentage of residents who rate water services as excellent, which are driven by individual perceptions, are subjective measures.

Therefore, to analyze the level of performance measurement validity, this study uses both objective measures of core criteria in public services from agency reports, and managers'

perceptions of criteria in public services based on the diverse forms of proximity that overviewed earlier in chapter 3: goal, informational, psychological, and social proximity. Objective measures refer to archival-based measures derived from content analysis of value content in each agency's annual performance and accountability report. Subjective measures refer to managers' perceptions of the values reflected in their agency's performance measurement criteria, as expressed in manager surveys. The difference between these objective and subjective measures is referred to as the 'gap'. Since the gap is computed as the difference between measurement validity in the agency's report and managers' perception, this gap will increase or decrease in a predictable way with value content—efficiency, service quality, customer service satisfaction, and social equity. The following hypotheses will be tested:

H4. Managers' objective proximity to the performance management process will increase the gap between archival-based and perceptual measures of value content.

H4a. Objective goal proximity will increase the gap between archival-based and perceptual measures of efficiency.

H4b. Objective goal proximity will increase the gap between archival-based and perceptual measures of service quality.

H4c. Objective goal proximity will increase the gap between archival-based and perceptual measures of customer service satisfaction.

H4d. Objective goal proximity will increase the gap between archival-based and perceptual measures of social equity.

H5. Managers' diverse forms of subjective proximity to the performance management process will increase the gap between archival-based and perceptual measures of value content.

H5a. Subjective goal proximity will increase the gap between archival-based and perceptual measures of value content.

H5a-1. Subjective goal proximity will increase the gap between archival-based and perceptual measures of efficiency.

H5a-2. Subjective goal proximity will increase the gap between archival-based and perceptual measures of service quality.

H5a-3. Subjective goal proximity will increase the gap between archival-based and perceptual measures of customer service satisfaction.

H5a-4. Subjective goal proximity will increase the gap between archival-based and perceptual measures of social equity.

H5b. Informational proximity will increase the gap between archival-based and perceptual measures of value content.

H5b-1. Informational proximity will increase the gap between archival-based and perceptual measures of efficiency.

H5b-2. Informational proximity will increase the gap between archival-based and perceptual measures of service quality.

H5b-3. Informational proximity will increase the gap between archival-based and perceptual measures of customer service satisfaction.

H5b-4. Informational proximity will increase the gap between archival-based and perceptual measures of social equity.

H5c. Psychological proximity will increase the gap between archival-based and perceptual measures of value content.

H5c-1. Psychological proximity will increase the gap between archival-based and perceptual measures of efficiency.

H5c-2. Psychological proximity will increase the gap between archival-based and perceptual measures of service quality.

H5c-3. Psychological proximity will increase the gap between archival-based and perceptual measures of customer service satisfaction.

H5c-4. Psychological proximity will increase the gap between archival-based and perceptual measures of social equity.

H5d. Social proximity will increase the gap between archival-based and perceptual measures of value content.

H5d-1. Social proximity will increase the gap between archival-based and perceptual measures of efficiency.

H5d-2. Social proximity will increase the gap between archival-based and perceptual measures of service quality.

H5d-3. Social proximity will increase the gap between archival-based and perceptual measures of customer service satisfaction.

H5d-4. Social proximity will increase the gap between archival-based and perceptual measures of social equity.

Organizational-Level Effects of Proximity and Performance

Since individual-level perceptions are nested in aggregate agency-level perceptions in systems of hierarchical governance, many studies tend to examine phenomena at both the individual and organizational levels (de Bruijn & Dicke, 2006). Organizational-level factors include organizational culture, human capital and capacity, agency support, leadership, red tape, and so forth (Brewer & Seldon, 2000). A number of studies have shown that organizational factors affecting individual and organizational performance (e.g., Osborne & Gaebler, 1992; Rainey & Steinbauer, 1999). Since several researchers have further emphasized that, organizations and individuals are interrelated because individual performance contributes to organizational performance (Brewer & Seldon, 2000; Pfeffer & Salancik, 1978).

With the importance of organizational-level effects, in Public Administration, many studies have conducted multilevel (ML) models including both individual- and organizational-level factors. For example, Jung and Lee (2013) use multilevel linear models to examine the relationship among goal properties, strategic planning capacity, and subjective organizational performance, considering within- and between-group variability. Hsieh (2013) also employed the multilevel method to examine job satisfaction, comparing multiple regression models. Feldman

(2002) used multilevel analysis to investigate why managers are likely to increase their work hours. Individual-level factors included demographic status and personality, and organizational-level factors were the organization's norms, leadership, and culture. Kobarg, Wollersheim, Welpe and Spörrle (2017) employed a multilevel method to examine individual ambidexterity (e.g. the individual's balanced pursuit of explorative activities). They found that individual ambidexterity positively affects individual performance in the public sector. At the same time, they found a negative relationship between supervisory ratio and performance in the public sector. Furthermore, many education studies have used this multilevel method to observe student achievement by inspecting the relationship between student learning and school climates, such as an instructional program or high levels of academic emphasis. Goddard, Sweeland, and Hoy (2000) conducted multilevel analysis to explore the school's effectiveness, while Duyar, Gumus, and Bellibas (2013) investigated teachers' self-efficacy and job satisfaction by including the relationship between administrative leadership and professional collaboration of teachers at the individual level and school characteristics at the organizational level. They found that principal leadership and collaborative practices had a positive influence on teachers' self-efficacy and job satisfaction within and across schools. These recent studies involved multilevel analysis and emphasized the importance of considering hierarchical characteristics in studies examining organizational structures because governments involve horizontal and vertical lines of authority and multiple programs within and across groups. Since the performance measurement system involves both organizational and individual level phenomena (Bloom, Hill, & Riccio, 2003; DeNisi, 2000; Lynn, Heinrich, & Hill, 2001; O'Toole & Meier, 2015), the multilevel method is used to examine among agency variability of proximity and value content in this dissertation.

CHAPTER 4

RESEARCH DESIGN AND METHODS

This chapter presents the data, variables, research model, and analytical methods for testing the hypotheses proposed in the previous chapters. Data from two different sources—Survey of Federal Managers on Organizational Performance and Management Issues (FMOPMI) in 2013 and 2017, and Performance and Accountability Reports (PARs) in 2013 and 2017—are utilized to estimate the level of performance measurement validity. As explained in Chapter 3, this study examines the congruence between value content taken from archival-based sources (agency annual performance and accountability reports) and managers' perceptions of agency intentions regarding value content, and thus assesses the level of measurement validity between objective and subjective measures. Two approaches are used: One approach involves aggregating managers' perceptions of value content at the agency level, which is a subjective, perceptual measure; and another approach is to perform content analysis on federal agency PARs to assess their value content, which is a more objective, archival-based measure.[3] To examine differences between managers' subjective perceptions of the values expressed in their agencies' performance measurement criteria, compared with objective measures derived from content analysis of the agencies' annual performance and accountability reports, the chapter also introduces all variables

[3] Objective measures refer to archival-based measures derived from content analysis of value content in each agency's annual performance and accountability reports. Subjective measures refer to manager's perceptions of the values reflected in their agency's performance measurement criteria, as expressed in manager surveys.

included in the analyses and explains how they are operationalized with examples. Then, the chapter discusses methodologies applied in this study. The research steps are summarized below:

1. Text analysis of PARs to discern their value content using both manual methods and ATLAS.ti qualitative analysis software to establish measurement reliability, and a radar chart is used to compare each value to check overall distribution differences of value content across stages—process, output, and outcome;
2. From 2013 and 2017 FMOPMI surveys, compute managers' perceptions of performance measurement validity in their agencies and run T-tests to identify statistically significant differences between the two surveys;
3. Compute the gap between value content at the agency level (from step 1 above) and managers' perceptions of value content expressed in the 2013 and 2017 surveys and aggregated up to the agency level (from step 2, above);
4. Compute the gap at the individual level for each value (defined as the difference between an employee's agency score and the employee's individual perception of performance measurement validity on the value);
5. Identify the factors explaining the gap with Multi-Level (ML) Modeling.

Data

Survey of Federal Managers on Organizational Performance and Management Issues (FMOPMI)

The first data source for this study is the 2013 and 2017 Survey of Federal Managers on Organizational Performance and Management Issues (FMOPMI). This web-based survey has been conducted regularly by the U.S. General Accountability Office (GAO) since 1997. The survey examines the perceptions of managers in the federal government on performance management issues, especially performance measurement concerns. It also includes questions

asking managers about their perceptions of various performance management activities including the Government Performance and Results Act of 1993 (GPRA) and the GPRA Modernization Act of 2010 (GPRAMA), such as perceptions of their agency's quarterly performance review activities, the presence and use of performance measures, measurement hindrances, performance information, and agency climate. FMOPMI enables the tracking of progress toward an agency's goal and helps managers to make resource-allocation decisions and adjust their goals for the following year.

For most questionnaire items, responses were coded 1 as 'to no extent,' 2 as 'a small extent,' 3 as 'a moderate extent,' 4 as 'a great extent,' and 5 as 'a very great extent.' The surveys went to a stratified random sample of mid-level and upper-level managers and supervisors in each agency and distinguished whether the manager or supervisor was a member of the Senior Executive Service (SES) or was in one of the "non-SES" agencies of the Chief Financial Officers (CFO) Act. In 2013, the observations across all agencies were 2,762 managers, which was about 69% of a stratified random sample of 4,391 managers from a population of about 148,300 mid-level and upper-level civilian managers working in the 24 executive agencies and government-wide organizations (GAO, 2014). In 2017, the response observations across agencies were 3,114 managers, which was about 67% of a stratified random sample of 4,395 managers from a population of about 153,779 mid-level and upper-level civilian managers and supervisors in the 24 executive branch agencies (GAO, 2017). The U.S. federal agencies in this study include 23 government agencies because the Department of the State and U.S. Agency for International Development shared a Joint Strategic Goal Framework. In this study, government agencies comprise the Department of Agriculture (USDA), Department of Commerce (DOC), Department of Defense (DOD), Department of Education (ED), Department of Energy (DOE), Department

of Health and Human Service (HHS), Department of Homeland Security (DHS), Department of Housing and Urban Development (HUD), Department of the State (DOS), Department of Justice (DOJ), Department of Labor (DOL), Department of Interior (DOI), Department of Transportation (DOT), Department of Treasury (USDT), Department of Veterans Affairs (DOV), Environmental Protection Agency (EPA), General Services Administration (GSA), National Aeronautics and Space Administration (NASA), National Science Foundation (NSF), Office of Personnel Management (OPM), Small Business Administration (SBA), Social Security Administration (SSA), and Nuclear Regulatory Commission (NRC). Appendix 1 lists the U. S. federal agencies in this study and their status as an executive or independent agency.

Performance and Accountability Reports (PARs)

The second data source was the 2013 and 2017 Performance and Accountability Reports (PARs) of federal agencies. The PARs are produced annually to share the agencies' progress toward achieving strategic goals and objectives. The PARs also present the agencies' financial statements, management challenges, and plans to overcome the challenges according to the Government Performance and Results Act of 1993, the Chief Financial Officers Act of 1990, and the Federal Financial Management Improvement Act of 1996. The GPRA Modernization Act requires agencies to develop an agency mission statement, objectives, and strategic measures to evaluate the agency's success in achieving the targets.

Main Variables of Interest

Using both the FMOPMI survey and the PARs, the purpose of this study is to explore the differences between the perceptions of performance measurement validity regarding the value

content of performance indicators in 2013 and 2017 and the value content expressed in the agency's annual report across the stages of a federal agency's performance (i.e., process, outputs and outcomes) in 2013 and 2017. The value includes efficiency, service quality, customer service satisfaction, and social equity. The main dependent variable is the performance measurement validity of value content, which is defined as the gap between archival-based measures of value content at the agency level that are more objective, and the managers' perceptions of value content at the individual level that are more subjective. The main independent variables are objective and subjective proximity. Additional individual-level variables are manager's accountability and external stakeholder attention. Control variables include SES and year at the individual level and institutional type and agency size at the organizational level. These variables are described in more detail below.

Values Explored in the FMOPMI Survey

Managers' perceptions of performance measurement validity of value content are expressed in the 2013 and 2017 FMOPMI surveys. Four values are named in the FMOPMI survey: efficiency, service quality, customer service satisfaction, and social equity. These values are highly relevant to federal agency performance as explained below.

Many scholars have debated the contribution of administrative science and efficiency in public administration (e.g., Dahl, 1947; Simon, 1946; Waldo, 1948). *Efficiency*, which can be described by a level of performance using the least input to achieve the highest output (e.g., Ostroff & Schmitt, 1993), still plays an important role in decisions to allocate resources for public services. Moreover, increasing the efficiency of services has remained high on the agenda of New Public Management in public sector theory and practice (Talbot, 1999; Brewer 2000, 2001). In this

regard, efficiency is measured by the survey item “we have performance measures that tell us if we are operating efficiently” in the 2013 and 2017 FMOPMI surveys.

Folz (2004, p. 213) explained that *service quality* has become an important element for all organizations. Therefore, many scholars pay attention to the value of service quality, advocating for Total Quality Management (TQM) in the public sector (Hu et al., 2009; Lewis, 1995; Milakovich, 1995; Swiss, 1992). Contemporary public managers provide public service based on the public’s interest and their needs. Public expects that the service delivery is completed on time with professional standards to assure service accuracy. Hence, service quality is assessed by the item “we have performance measures that tell us about the quality of the products or services we provide” in the 2013 and 2017 FMOPMI surveys.

Citizen/customer service satisfaction has been defined as an overall assessment of the services that can be explained by the likelihood of returning to the service in the future (McDougall & Levesque, 2000). Customer satisfaction also refers to the “customer’s response to the evaluation of the perceived discrepancy between prior expectation and the actual performance of the product as perceived after its consumption” (Tse & Wilton, 1988, p. 204). Customer service satisfaction is evaluated by the item “we have performance measures that tell us whether or not we are satisfying our customers” in the 2013 and 2017 FMOPMI surveys.

Regarding *social equity*, Selden, Brewer, and Brudney (1999) pointed out that equity and fairness in the distribution of public goods play a critical role in service delivery. All levels of government focus on "fairness in the distribution of goods and services among the people in an economy" (Friedmann, 2002, p. 58). However, scholars have pointed out that under the GPRA, little attention was given to social equity (e.g., Frederickson, 2010). In 2002, the Standing Panel on Social Equity in Governance of the National Academy of Public Administration

acknowledged that there are important linkages between social equity and the contemporary performance measurement movement. Equity is associated with "how well public organizations are able to tailor service provision to meet the needs of the diverse groups of citizens that they serve" (Andrews & Van de Walle, 2013, p. 13). Hence, social equity is gauged by the survey question, "we have performance measures that tell us how equitably our products or services are distributed among our customers, as appropriate" in the 2013 and 2017 FMOPMI surveys.

Measuring Value Content in PARs

To explore managers' perceptions of performance measurement validity, the four values explored in this study are efficiency, service quality, customer service satisfaction, and social equity. The researcher run software and manually counted the number of times each value was expressed in the PARs across each stage of performance—process, outputs, and outcomes. Counting the data is the most commonly employed method in qualitative research, promoting scientific rigor and objectivity (Porter 1995). Some scholars have expressed concern that content analysis including interviews is sometimes used for inappropriate generalization to a larger population or for making statistical inferences. Under the right circumstances, creating and analyzing counts based on systematic reviews can, however, be a productive and distinctive strategy (Kimberly & Kramer, 2015; Morgan, 1993). Qualitative data that contain extensive descriptions and nuanced understandings help to identify complex patterns that might not be detected from an established quantitative version (Langley, 1999). Counting the frequency of the codes helps to detect patterns that may contain multi ideas which cannot be detected in quantified data (Langley, 1999; Morgan, 1993). Thus, the quantification of qualitative data has emerged as an analytic strategy in recent years (e.g., Onwuegbuzie et al., 2009). Since this study

tracks trends of the value content of agency performance measurement criteria based on the stage of performance—process, outputs, and outcomes and compares them across agencies, the counting technique was considered appropriate.

Values Explored in the PARs

Andrews and Entwistle (2013) specify four main dimensions of public service efficiency: productive, allocative, distributive, and dynamic efficiency. According to Andrews and Entwistle (2013), productive efficiency refers to the relative inputs required to achieve the basic outputs of production. This implies an effort to maximize outputs and minimize inputs (Farrell, 1957). Allocative efficiency is the match between the demand for service and supply in order to provide for a better democracy while distributive efficiency can be defined as the distribution of resources or services between citizens and the relative cost to government of that distribution—for example, a Pareto-efficient distribution of resources. Dynamic efficiency focuses on the allocation of resources between current and future consumption. The overall efficiency of public service provision can be gauged by the extent that the government includes the ratios indicating the financial cost of producing output or outcomes (Andrews & Entwistle, 2013). Efficiency is captured in the PARs based on Andrews and Entwistle's demonstration of efficiency. Efficiency was coded when performance indicators contained any of these words: *efficiency, efficient, cost per, cost-saving, cost-effective, reduce/decrease expense, reduce/decrease waste, program-operating efficiency, discount, gross margin, cost avoidance, reduce expense, the value of costs, reduce vacant space,* or *unit cost.*

Attempts to define and measure service quality to everyone's complete satisfaction are analogous to the challenge of 'herding cats.' Many studies have, however, endeavored to define

and measure it. Service quality in the public sector has been described as having two dimensions: professional quality (e.g., Donabedian, 1989) and management quality (e.g., Curry & Herbert, 1998). Professional quality relates to the "processes and techniques used to meet customer needs through organizational audits and setting standards" while management quality relates to "the efficient use of resources to meet customer needs" (Agus et al., 2007, p. 180). The conceptual model for assessing service quality is comprised of several major determinants: access, communication, competence, courtesy, credibility, reliability, responsiveness, security, tangibility, and understanding the customer (Agus et al., 2007). Reliability concerns whether the outcome of service delivery is as promised, and it can be achieved by training employees with ample information and by establishing standards in order to assure service accuracy. On the other hand, responsiveness is determined by whether the service delivery is completed on time (e.g., timeliness) (Hu et al., 2009). Therefore, service quality is captured by certifying whether the service is delivered by trained employees according to the standards and principles for improving competence, credibility, and security as professional quality and by ensuring whether the service is distributed for completing responsiveness and reliability as management quality. Accordingly, service quality was coded when performance indicators contained any of the following words: *credibility, credible, responsible, responsive, responsibility, qualified, high-quality, improper, error, reliable, reliability, certified, expert, credibility, align with the mission, professional, skilled workforce, accuracy, accurate, desired condition, good condition, on time, timely, audit, training, monitor, monitoring, principle, guideline, standard, rule, law, compliance, modernize, scientific data, well-being, systematic, programmatic assessment, strategic, evidence-based, responsive, responsible, competency, consistent with law, best value,* or *expert review*.

Customer service satisfaction is interpreted literally— if respondents agree that performance indicators or strategic goals are established with the intention of improving customer service satisfaction. Many scholars have used the terms citizen satisfaction and customer satisfaction interchangeably. Hatry et al. (1973) explain that customers are the end-users of public goods and services while the citizens are the actual beneficiaries and potential customers who pay the tax to support the services (Talbot, 1999). In this study, the terms customers, citizens, and clients were used. The customer service satisfaction value was measured by drawing upon performance indicators that contained any of the following words: *satisfaction, satisfied, clients rate, citizen-centered service, customer service, customer needs, customer value, rate services good to excellent, satisfactory, customer feedback, rating, customer loyalty, or customer trust.*

Social equity can be measured with horizontal and vertical perspectives by drawing upon indicators or agencies' goals that specify providing a fair distribution of public services and delivering public services to people with special needs. According to its definition in the context of public service provision, *social equity* consists of the elements 'horizontal' and 'vertical' (Rolle & Liu, 2007). Chitwood (1974) notes that horizontal equity refers to "the degree to which programs treat all eligible beneficiaries equally" (Langbein, 1980, p.21). Vertical equity is concerned with the redistribution of public services within a society based on factors such as geographic location, health, gender, age, income, etc. An example of vertical equity is progressive taxes. Therefore, this study explores equity value even though equity measures are often overlooked when measuring performance, partly because there is a strong tendency to (over) rely on efficiency measures (Radin, 2006). In this study, social equity was coded when indicators contained any of the following terms: *equal, equity, inequity, fair, unfair, access,*

accessibility, discrimination, low wage workers, equal employment opportunity, diversity and inclusion strategic plan, vulnerable people, developing countries, immigration, migrants, undocumented, underrepresented group, underserved group, refugee, vulnerable population, Natives, Black, tribal, tribe, underage, child, children, women, African-American, Indian, disabled, disability, handicap, diversity, racial, ethnic minority, elderly, homeless, poverty, low-income, or *multi-family.*[4]

Archival analysis of PARs allowed the researcher to manually count the number of performance indicators for each value and in each stage of the performance management process. To improve validity and reliability, ATLAS.ti qualitative analysis software was utilized to check for the terms mentioned above. Pearson correlation analysis was conducted between the manual counts and electronic counts. The average Pearson correlation, shown in Table 4.1, was relatively high at .89, and ranged from 0.77 to 0.97 across the value categories—process, output, and outcome stages. Differences were resolved by a second round of manual checking in which each individual difference was scrutinized, and the researcher made a final decision.

Table 4.1 Pearson Correlations Between Manual and Software Analysis of Value Content Across Stages of the Policy Process

	Outcome	Output	Process	Average
2013	0.96	0.94	0.77	0.89
2017	0.97	0.87	0.83	0.89

Collective Managers' Perceptions of Performance Measurement Validity

[4] Several examples of how to distinguish value content are shown in Appendix 2.

Some social science studies have aggregated individual-level data to higher levels, such as the organizational, institutional and team levels (Cole, Bedeian, Hirschfeld & Vogel, 2010; Kimberly, 1980). Before aggregating individual perceptions of performance measurement validity by exploring value content, this study notes that many scholars have suggested conducting internal consistency tests to gain a rigorous estimate of the extent to which the aggregated responses with sufficient within-group agreement have been presented in various forms, such as the Bland-Altman plot, *t*-test, Pearson correlation coefficient, and Intraclass coefficient (Bland & Altman, 1986; Brown Jr., Lucero & Foss, 1962; Bruton, Conway, & Holgate, 2000; Hopkins, 2000). Some were not ideal measures of reliability for this study. The Bland-Altman plot and *t*-test were used for analyzing agreement, while the Pearson correlation coefficient was used for strength of correlation. A more desirable measure of reliability includes both the degree of correlation and agreement between measurements (Koo & Li, 2016). The Intraclass Correlation Coefficient (ICC), which includes the correlation and agreement between measurements, refers to the "measure of the reliability of the number of different raters to measure subject similarity" (Bobak, Barr, & O'Malley, 2018, p.94). The concept of ICC was introduced and developed by many scholars (e.g., Glick & Roberts, 1984; McGraw & Wong, 1996). Since then, the ICC has become an important measurement instrument and the most common index in reliability studies. The ICC index has been used to justify the aggregation of data (e.g., Andrews & Entwistle, 2013; Jung, 2012; Moynihan & Pandey, 2004). ICC (1) is an index of group variability, and ICC (2) is an index of the reliability of group means (Bliese, 2000). Therefore, ICC (1) compares the proportion of variance among agencies to the variance within agencies, while ICC (2) is an estimation of the reliability of the group means,

differentiating them from one another (Raudenbush & Bryk, 2001).[5] Researchers can interpret the respective magnitudes of ICC (1) and ICC (2) as "the level of observed variance of single score and average score that is affected by clustering" (Shieh, 2016, p. 995). For example, one study explored the relationship between the influence of classroom climate perception and individual students' levels of academic achievement with a two-level analysis. If the results of ICC (1) value was 0.4 and ICC (2) value was 0.6, the researcher could infer that an ICC (1) value of 0.4 means that 40% of the observed variance in students' achievement scores could be explained by between-classroom differences compared to the total variance in achievement scores. An ICC (2) value of 0.6 indicates that 60% of the observed total variance in classroom average scores occurred at the classroom level. Therefore, the researcher could make inferences about "the reliability of single score and average score, respectively" by conducting ICC (1) and ICC (2) (Shieh, 2016, p. 995).

In this study, the ICC estimates and their 95% confidence intervals were calculated using the statistical package STATA 13 based on a one-way random-effects model presented in Table 4.2. The ICC (1) ranged between 0.04 and 0.27. This suggests that around 4%-27% of the variance in perceptions of value content can be explained by between-agency differences compared to the total variance in performance measurement validity. Regarding ICC (2), as a common rule of thumb, higher positive scores close to 1 indicate a high degree of agreement among respondents. Based on the guideline from Koo and Li (2016), ICC (2) is a value between 0 and 1, where values below 0.5 indicate poor reliability, between 0.5 and 0.75 denote moderate reliability, between 0.75 and 0.9 suggest good reliability, and greater than 0.90 indicate excellent

[5] ICC (1) $= \frac{MSB-MSW}{MSB+(K-1)MSW}$, and ICC (2) $= \frac{MSB-MSW}{MSB}$, where MSB is the between-group mean square and MSW is the within-group mean square (Shieh, 2016).

reliability. The average of the ICC (2) for value content ranged from 0.71 to 0.84 in the 2013 and 2017 survey data. The ICC statistics denote good consistency in responses among individuals of the same agency in the 2013 and 2017 survey data.

Table 4.2 Degree of Consistency of Aggregated Statistics at the Agency Level

Year	Variable	F	ICC (1)	ICC (2)
2013	Efficiency	5.30***	0.0730	0.81
	Service Quality	5.62***	0.2423	0.82
	Customer service satisfaction	6.38***	0.0532	0.84
	Social Equity	3.47***	0.0827	0.71
2017	Efficiency	4.41***	0.0450	0.77
	Service Quality	4.34***	0.2741	0.77
	Customer service satisfaction	5.27***	0.0463	0.81

* N=23 agencies

ICC (1) compares the proportion of variance among agencies while ICC (2) is an estimation of the reliability. Based on the ICC (1) and the ICC (2) values, these variables justify the aggregation of individual perceptions of value content at the agency level. Several ways of aggregating responses to an upper level have been adopted (Klein, Conn, Smith & Sorra, 2001; van Mierlo et al., 2009). First, the direct consensus model is the most familiar and widely used in multilevel research (Cole et al., 2010). It computes "within-group agreement of scores to index consensus at the lower-level and to justify aggregation of lower-level scores to represent scores at the higher level" (Chan, 1998, p.237). The mean-based approach considers that employees have built an agreement of organizational collective cutoff values from the aggregation for the individual scores (Chan, 1998). According to this view, previous studies have found that it is

more suitable to aggregate the individual-level data to the organizational level than to the department level (Chan, 1998). Second, a referent-shift consensus model similar to direct consensus composition is also broadly used for within-group consensus, as indexed by agreement of lower-level attributes. The difference between the referent-shift consensus and direct consensus model is that the lower-level attributes originate from the original individual-level construct in the referent-shift consensus model. The change in referent leads to different results with a new form of the original focal construct. Third, the additive model uses simple summing or averaging lower-level units to present the value regardless of the variance among individuals (Chan, 1998). According to this view, there are questions about measurement accuracy, including random error and sources of bias (Glick, 1985). Fourth, the dispersion-composition model refers to individual differences within the group. Finally, process composition is concerned with the mechanism and process from the lower level of conceptualization to the higher level by identifying critical parameters and their interrelationships. This composition model is more interested in 'proceduralization' from episodes or changes in behaviors (Anderson, 1982).

Specifying a satisfactory composition model directly associated with how the construct is operationalized is a critical component of research. This study should not use the referent-shift consensus model because it is not directly composed from an individual level that needs sufficient agreement among group members. Researchers need to be careful about using the additive model, as a simple summation of individuals' perceptions raises accuracy issues concerning variance among individuals. The use of the dispersion-composition model is limited in that it needs an empirical prerequisite of multimodality in the within-group distribution for aggregating the lower-level to the higher-level construct. Without a concrete practical algorithm

to compose the lower-level process to the higher-level process, process composition is inapplicable for measuring outcome and output. Consequently, this study uses the mean-based approach with the average score of individual perceptions of the validity of performance measures.

Independent Variables

In order to examine the relationship between the validity of performance measures and the managers' proximity to the performance measurement process, this study uses two forms of proximity: objective and subjective. By the definition, objective proximity is measured with one item that emphasizes the managers' actual involvement in performance measures while subjective proximity includes social, informational, organizational, and goal proximity. Principal Component Analysis (PCA) was used to measure managers' subjective proximity at the individual level. In this section, the following items are used to measure individual-level perceptions of performance management activities and create different types of proximity effects using PCA. In the PCA with varimax rotation, results are obtained for five factors: objective goal, subjective goal, informational, psychological, and social proximity. Cronbach's alpha, which assesses the degree of reliability of the measuring tool, ranged from .67 to .88, which exceeded the required threshold of >.70 (Nunnally, 1978), as shown in Table 4.3.

The measurement items for each factor are:

- Objective goal proximity

- I have been involved in creating the cross-agency priority goals. (*Goal-O1*)
- I have been involved in creating my agency's priority goals. (*Goal-O2*)

- Subjective goal proximity

  - How familiar are you with one or more of your agency's priority goals? (*Goal-S1*)
  - How familiar are you with one or more of the cross-agency priority goals? (*Goal-S2*)

- Informational proximity

  - My agency's performance information is easily accessible to managers at my level. (*Info-S1*)
  - My agency's performance information is available in a format that is easy to use. (*Info-S2*)
  - I have access to the performance information I need to manage the program(s) /operation(s) /project(s) that I am involved with. (*Info-S3*)
  - Performance information is available in time to manage the program(s) /operation(s) /project(s) that I am involved with. (*Info-S4*)

- Psychological proximity[6]

  - Difficulty determining meaningful measures. (*Psy-S1*)
  - Difficulty obtaining valid or reliable data. (*Psy-S2*)
  - Difficulty obtaining data in time to be useful. (*Psy-S3*)

- Social proximity

  - Employees in my agency receive positive recognition for helping the agency accomplish its strategic goals. (*Soc-S1*)
  - Agency managers/supervisors at my level effectively communicate performance information on a routine basis. (*Soc-S2*)

[6] Psychological proximity variables are revere coded.

- Agency managers/supervisors at my level use performance information to share effective program approaches with others. (*Soc-S3*)

Additional Independent Variables

Besides proximity effects, this study includes two additional independent variables based on previous research findings discussed in Chapter 2: manager's accountability and external stakeholder attention. Manager's accountability and external stakeholder attention were factor scores derived from the PCA with Varimax rotation. Cronbach's alpha ranged from .83 to .92, exceeding the required threshold of >.70 (Nunnally, 1978), as shown in Table 4.3. The measurement items for each variable are shown below:

- Manager's Accountability

- Agency managers/supervisors at my level are held accountable for agency accomplishment of its strategic goals. (*Acc-1*)
- Agency managers/supervisors at my level are held accountable for the results of the program(s)/operation(s)/project(s) they are responsible for. (*Acc-2*)

- External Stakeholder Attention

- Congressional committees pay attention to agency's use of performance information in management decision making. (*Ex-1*)
- The audits community (e.g., GAO, Inspectors General) pays attention to agency's use of performance information in management decision making. (*Ex-2*)

Control variables

This study uses control variables to improve model specification. As described in Chapter 2, control variables for the individual level are SES status and year, and organizational level control variables include agency size and institutional type due to limited data. The measurement items for each control variable are shown below:

- Senior Executive Service (SES)
  - SES / non-SES (binary)
- Year
  - 2013
  - 2017
- Agency Size
  - The number of employees (divided by 1000)
- Institutional Type[7]
  - Executive Department / Independent Agency (binary)

Table 4.3 Summary of Principal Component Factor Analysis by Year (Varimax Method)

Variable	**Cronbach α**	
	2013	2017
Objective Goal Proximity	0.67	0.72
Subjective Goal Proximity	0.71	0.73
Informational Proximity	0.83	0.87
Psychological Proximity	0.85	0.85
Social Proximity	0.82	0.81
Manager's Accountability	0.85	0.88
Stakeholder Attention	0.83	0.84

[7] Institutional type includes executive departments and independent agencies. See Appendix 1 for details.

Methods

This study examines the effects of various forms of proximity on performance measurement validity concerning value content on four different values (efficiency, service quality, customer service satisfaction, and social equity) and in three stages: process, output and outcome. Two separate sources of data were used: (1) managers' collective perceptions of value content taken from the 2013 and 2017 FMOPMI surveys, representing subjective and perceptual measures; and (2) Measures taken at the agency level from the PARs, representing objective and archival-based measures.

Several techniques were used to test the hypotheses proposed in the previous chapters. First, archival analysis of the agency's annual performance and accountability reports determined their value content by using manual methods and ATLAS.ti qualitative analysis software, and resolving differences by rechecking. A radar chart is used since it is known as a good analytical tool for evaluating organizational performance (Bogan & English, 1994).[8] The radar chart has four or more axes integrated into a single radial figure and presents the cases in the organizational performance criteria simultaneously showing changes over time (Sezhian et al., 2011). Thus, it employed for analyzing how several criteria are used across stages—process, output, and outcome simultaneously over the years. Second, from the 2013 and 2017 FMOPMI surveys, managers' perceptions of performance measurement validity in their agencies were aggregated to the organizational level by computing a mean value for each agency and using a

[8] A radar chart is "an increasingly popular way to present spatial data in a visually interesting format" (Feldman, 2013). It has been used for analyzing physical performance and health (Lin et al., 2018), key successful factors for building relationships between mangers and customers (Rajesh et al., 2010), comparative policy stances among emerging market economies countries, and comparing research productivity in industrial health among countries (Scutaru et al., 2010).

means-based approach (Chan, 1998). The repeated *t*-test for the time variable was conducted to identify significant differences between 2013 and 2017. A repeated measurement test assesses the differences in between-subjects over time (Hedeker & Gibbons, 2006; Huta, 2014). Third, this study computed the gap between an agency's value content expressed in the 2013 and 2017 PARs and aggregated managers' perceptions of performance measurement validity (both aggregated to the agency-level). Fourth, this study computed the gap between value content expressed in the 2013 and 2017 PARs and individual managers' perceptions of value content expressed in the 2013 and 2017 surveys (both disaggregated to the individual level in which the manager's agency score was considered the manager's individual score). Lastly, the factors explaining managers' perceptions of performance measurement validity at the individual level were identified and examined using Multi-Level (ML) modeling (the main variables of interest were objective and subjective proximity).

This study conducted ML modeling characterized by a hierarchical structure since the data included individual- and organizational-level measures. Hierarchical levels of grouped data are commonly used in many fields, especially organizational research (e.g., Bryk & Raudenbush, 1989). Lynn and Heinrich (2000) also employed ML modeling because governance is a complex phenomenon that cannot be understood by examining a single organization. This method investigates the relationship within and between levels of data, explaining the variance among variables at different levels (Mass & Hox, 2005).

Moreover, regarding ICC (1), Hox (1998) suggested 5% as a threshold for using multilevel factor models. In this study, ICC (1) showed around 4%-27% of the variance in individual perceptions of value content that is explained by group membership. Huta (2014) explains that adequate group sample sizes need to be 15 or greater in social science studies and

the number of agencies in this study is 23 each year.[9] By fulfilling the overall conditions for employing multilevel factor models, therefore, the ML modeling is the preferred method to estimate performance measurement validity of value content nested at the agency level under a system of hierarchical governance. All equations were computed based on Hamilton's (2012) guide.

In the fixed model, the equation was written as:

$$y_i = \beta_0 + \beta_1 \beta X_{1i} + \beta_2 \beta X_{2i} + \beta_3 \beta X_{3i} + \beta_4 \beta X_{4i} + \beta_5 \beta X_{5i} + \epsilon_i$$

This study included not only a set of β coefficients that described all the proximity effects, but also a random intercept u_0, which varied from agency to agency. In the random intercept model, the equation for case *ith* individuals within unit *jth* agency was written as:

$$y_{ij} = \beta_0 + \beta_1 \beta X_{1ij} + \beta_2 \beta X_{2ij} + \beta_3 \beta X_{3ij} + \beta_4 \beta X_{4ij} + \beta_5 \beta X_{5ij} + u_{0j} + \epsilon_{ij}$$

This study included the random intercept and slope model, allowing intercepts and slopes to vary across agencies. In the random intercept and slope model, the equation including random slopes intercept u_{1j} on predictor x_1, and random intercepts u_{0j} for each of j agencies was written as:

$$y_{ij} = \beta_0 + \beta_1 \beta X_{1ij} + \beta_2 \beta X_{2ij} + \beta_3 \beta X_{3ij} + \beta_4 \beta X_{4ij} + \beta_5 \beta X_{5ij} + u_{0j} + u_{1j} X_{1ij} + \epsilon_{ij}$$

[9] Surveys of Federal Managers on Organizational Performance and Management Issues include 24 executive branch agencies, of which 23 are included in this study since the Department of State and the United States Agency for International Development shared a Joint Strategic Goal Framework.

Since random effects represent grouping variables, it allows the estimation of variance in the response variable among groups (Breslow & Clayton, 1993; Mass & Hox, 2005). Using random effects "reduces the probability of false positives (Type I error rates) and false negatives (Type II error rates)" (Grawley, 2013; Harrison et al., 2018, p. 2). Moreover, estimating the statistical level of variation among groups can be informative to understand the variation of groups. While the fixed effect assumes "the group means are all independent of one another and share a common residual variance," the random effects assume "the group means are only a subset of the realized possibility drawn from a population means that follow a normal distribution with its own mean (μ_{group}) and variance ($\sigma 2_{group}$)" (Harrison et al., 2018, p. 4). Therefore, inferring among group variation in means is more intuitive and mathematically sophisticated than fixed effects.

The statistical analyses and results for each hypothesis are presented in the next chapter along with some patterns of findings. These patterns will be reflected on in the concluding chapter.

CHAPTER 5

RESULTS: PERFORMANCE MEASUREMENT VALIDITY AND PROXIMITY

The purpose of this chapter is to examine the relationship between proximity and performance measurement validity of value content (i.e., efficiency, service quality, customer service satisfaction, and social equity), which is defined as the level of congruence between objective, archival-based measures taken from agency performance reports; and subjective, perceptual measures compiled from manager surveys. This chapter summarizes the findings with descriptive statistics, correlation matrixes, *t*-test results, a radar chart showing the variety of value content by stages of the policy process across years, and the results of empirical analyses from the Multi-Level (ML) model.

Subjective Measures Analyses Results

Descriptive Statistics

Managers' perception of their agency's value content was analyzed using the Survey of Federal Managers on Organizational Performance and Management Issues (FMOPMI). Table 5.1 shows the descriptive statistics of all variables presented in this dissertation. While efficiency and customer service satisfaction validity had the highest mean value, 3.45, social equity validity had the lowest mean value, 2.35, among the four types of value content. Regarding the main independent variables, most of the proximity variables ranged from 0 to 5 but subjective goal proximity (denoted Goal-S in Table 5.1) ranged from 0 to 6. There was no serious normality

problem based on the PP plot, residual distribution, and Variance Inflation Factors (VIFs) for multicollinearity, which ranged from 1.01 to 1.57. The average VIF score of efficiency, service quality, and customer service satisfaction was approximately 1.24 and the average VIF score of social equity was approximately 1.25. Heteroskedasticity was, however, detected. Thus, the analysis used the robust estimator option, which has a smaller standard error, to deal with the main assumptions of the Gauss-Markov theorem, homoskedasticity.

Table 5.1 Descriptive Statistics

Variable	Observations	Mean	SD	Min	Max
Gap of Efficiency	4288	3.22	1.23	-0.77	5.00
Gap of Service Quality	4284	1.86	1.53	-3.64	4.68
Gap of Customer Service Satisfaction	4284	3.25	1.25	-0.98	5.00
Gap of Social Equity	2200	1.73	1.71	-1.79	5.00
Objective Goal Proximity	2653	-0.06	1.07	-1.96	3.28
Subjective Goal Proximity	4714	4.36	1.10	0	6
Informational Proximity	2653	0.03	1.03	-4.33	2.61
Psychological Proximity	2653	-0.07	1.01	-2.77	2.81
Social Proximity	2653	0.06	0.99	-4.52	3.13
Manager's Accountability	4351	-0.01	1.01	-4.28	1.58
External Stakeholder Attention	4351	0.06	0.97	-1.55	2.42
Agency Size (/1000)	4788	10.15	1.40	7.11	12.74
Institutional Type	4788	0.67	0.47	0	1
SES	4788	0.20	0.40	0	1

Table 5.2 displays the correlation matrix for the variables. The highest correlation coefficient is 0.47 between social proximity and manager's accountability, and the next highest is 0.41 between informational proximity and the efficiency value. According to Cohen (1988), a correlation coefficient of 0.30-0.50 is considered a moderate correlation. Except for those two relationships, most correlation coefficients are 0.3 or under, meaning low-moderate correlations exist.

Table 5.2 Correlation between Variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Gap of Efficiency	1.000													
2. Gap of Service Quality	0.408	1.000												
3. Gap of Customer Service Satisfaction	0.539	0.532	1.000											
4. Gap of Social Equity	0.311	0.357	0.344	1.000										
5. Objective Goal Proximity	0.131	0.176	0.117	0.211	1.000									
6. Subjective Goal Proximity	0.105	0.101	0.057	0.025	0.312	1.000								
7. Informational Proximity	0.413	0.276	0.371	0.221	-0.02	0.142	1.000							
8. Psychological Proximity	0.232	0.103	0.187	0.102	-0.01	-0.02	0.038	1.000						
9. Social Proximity	0.215	0.190	0.215	0.145	-0.004	0.115	-0.05	0.021	1.000					
10. Manager's Accountability	0.259	0.162	0.223	0.105	0.026	0.132	0.228	0.070	0.479	1.000				
11. External Stakeholder Attention	0.114	0.093	0.087	0.164	0.219	0.140	0.094	-0.089	0.145	0.036	1.000			
12. Agency Size (/1000)	0.020	0.087	0.062	0.025	-0.01	-0.11	-0.03	-0.03	0.022	0.470	0.005	1.000		
13. Institutional Type	-0.012	0.016	0.025	-0.009	0.019	-0.168	-0.107	-0.023	-0.017	-0.070	-0.001	0.470	1.000	
14. SES	0.020	0.087	0.062	0.025	-0.01	-0.11	-0.03	-0.03	0.022	0.470	0.005	-0.018	0.007	1.000

Objective Measures Analyses Results

This section presents the results from archival-based measures using the 2013 and 2017 Performance and Accountability Reports (PARs). In PARs, the number of total indicators was 2,019 in 2013 and 1,976 in 2017 (see Tables 5.3 and 5.4). The number of efficiency and customer service satisfaction value measures were more frequently expressed in 2017 than in 2013, supporting hypotheses H1a and H1c. This means that there are more performance indicators, strategic objectives, and strategic goals on efficiency and customer service satisfaction in 2017 than 2013. Service quality measures were more frequently found in 2013 rather than 2017 (total n=624 and 515, respectively), and social equity measures were the second most often expressed in 2013 and 2017 (total n=248 and 235, respectively) across agencies. This means that there are more performance indicators, strategic objectives, and strategic goals on service quality and social equity in 2013 than 2017. Efficiency was the most frequently expressed value in the Department of Energy (DOE) and General Services Administration's (GSA) reports in 2013 and 2017, while service quality was the most frequently expressed value in the 2013 and 2017 reports of the Nuclear Regulatory Commission (NRC), Social Security Administration (SSA), and Department of Education (ED). Customer service satisfaction was the most frequently expressed value in the Department of Homeland Security (DHS) and Department of Veterans Affairs' (DOV) reports in 2013 and 2017, while social equity was the most frequently expressed value in the 2013 and 2017 reports for the Department of Health and Human Services (HHS) and SSA.

Performance measurement validity was measured as the size of the gap between archival-based measures, which are relatively objective, and managers' perceptions of value content aggregated to the federal agency level, which are relatively subjective. There are two ways of

analyzing value content across the stages of the policy process model shown in Table 5.3. First, considering the value itself within each stage of the policy process model, efficiency was more frequently expressed in the outcome stage, supporting H2a, and service quality was more frequently expressed in the outcome stage, supporting H2b. On the other hand, customer service satisfaction was more frequently expressed in the process stage than in the output and outcome stages in both 2013 and 2017, not supporting H2c. Social equity was more frequently expressed in the outcome stage than in the input and output stages in only 2017, partially supporting H2d.[10]

Table 5.3 Frequencies of Value Content in the Process, Output, and Outcome Stages

	Efficiency		Service Quality		Customer service satisfaction		Social Equity	
	2013	2017	2013	2017	2013	2017	2013	2017
Process	23 (1.54%)	24 (1.60%)	483 (32.33%)	389 (26.00%)	34 (2.34%)	52 (3.48%)	177 (11.85%)	171 (11.43%)
Output	22 (5.28%)	24 (6.61%)	103 (24.70%)	78 (21.49%)	2 (0.48%)	1 (0.28%)	54 (12.95%)	49 (13.50%)
Outcome	6 (5.56%)	8 (6.84%)	38 (35.19%)	48 (41.03%)	0 (0.00%)	2 (1.71%)	17 (15.74%)	15 (12.82%)
Total	51	56	624	515	36	55	248	235

Second, from the stage-oriented perspective, this result can be shown on a radar chart which is a graphical method of displaying multiple data points (Albouy et al. 2010). The reader can easily compare the relative frequencies of each value along its own axis, and overall differences are displayed by the general size and shape of the polygons (Sezhian, 2011).[11] Figure 5.1 displays the results for each value in each stage of the performance measurement process in

[10] Specific frequencies and the ratio (frequencies / total number of indicators) based on the stages are shown in Appendix 3.

[11] All axes begin at a central point designated 0. The scores for each variable are recorded in increments of 10 and range from 0 to100.

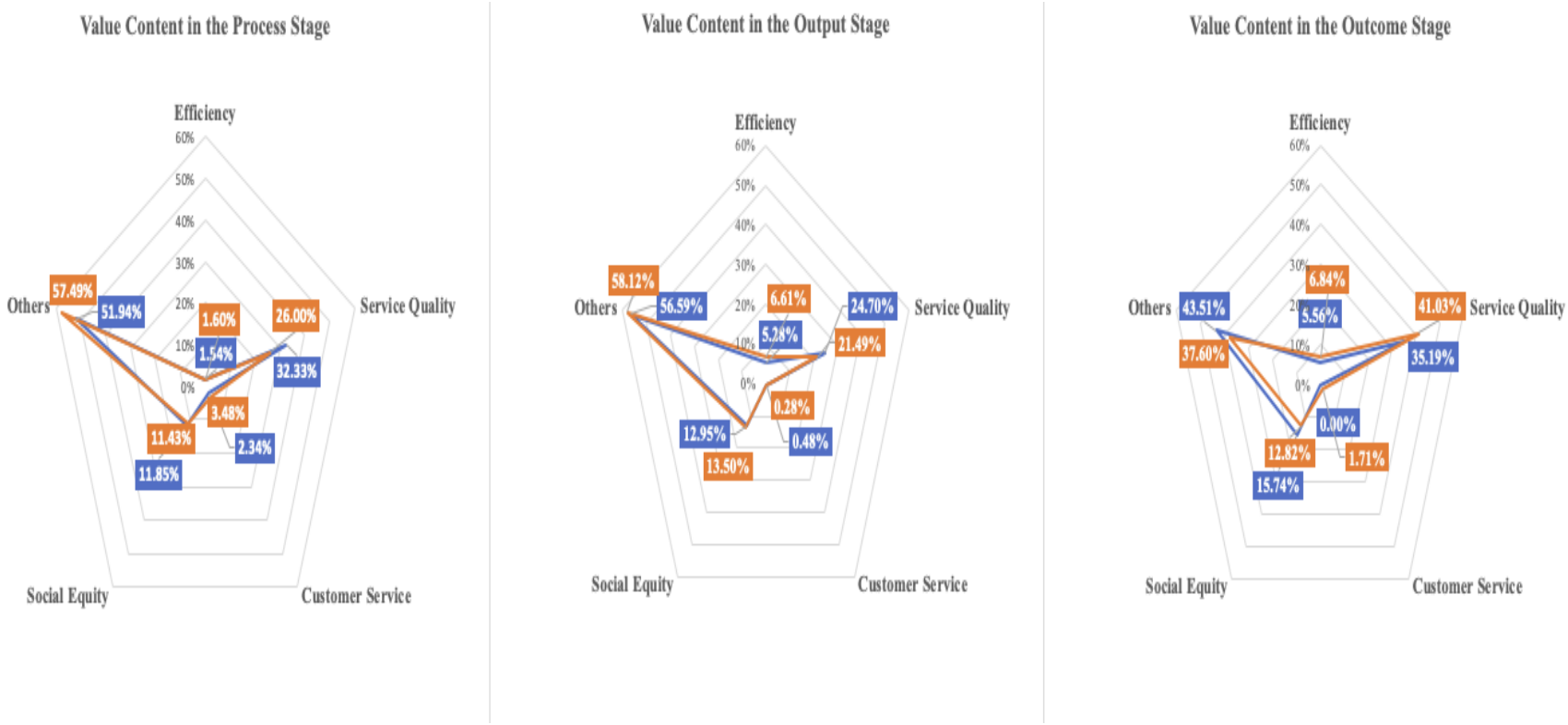
2013 and 2017. In each polygon, the blue lines represent value content in 2013 and the orange lines denote value content in 2017. Since the total number of performance indicators (process), strategic objectives (outputs), and strategic goals (outcomes) are different, each number was converted into a percentage in order to easily compare the three stages across years. One way to interpret the results of a radar chart is to focus on its balance or proportionality. A balanced or proportional chart is one in which the lines are roughly the same length and when superimposed on each other, produce a symmetrical result. In this case, the lines represent the number of performance measurement indicators devoted to each value. Accordingly, a disproportionate polygon indicates value imbalance. Another way to interpret the results of the radar chart is to focus on its filled area. Leary et al. (2002) have found that 'a well-performing unit is always likely to have a larger area than a less well-performing unit'' (p. 754). Feldman (2013) has, however, argued that a filled radar chart is not appropriate for comparing social indicators. The size of the area, calculated by multiplying length and width, is not suitable.[12] Hence, this study uses the first method of interpreting the length of the polygon.

In Figure 5.1, overall differences are apparent in the size and shape of the polygons. The most frequently presented value in performance measures is service quality, while the least presented values are customer service and efficiency across stages and over the years. In addition, according to the radar chart, the increasing disproportionality across the process, output and outcome stages of the performance measurement process is not significant There are only a few changes across stages indicating that value coherence exists between 2013 and 2017. Coherence also exists among the values in the degree to which they are organized coherently and

[12] This study used a pentagon shaped radar chart. This pentagon has a credibility problem. "As long as the two smaller pentagons are formed by connecting the midpoints of the next larger pentagon, the two ratios of the areas will always be equal, and the two ratios of the perimeters will always be equal" (Zbiek, 1996). This could yield biased information.

consistently across the process, output, and outcome stages of the policy process. New Public Management (NPM) reforms call for explicit use of standards and measures of performance based upon both Hood's (1991) and Moore's (1995) formulations of public value in the public sector. Even though there are different plans and schemes for managing performance in the Obama and Trump administrations, the U.S. federal government's consistent efforts are revealed in the radar chart, which displays value coherence.

Figure 5.1 Value Content Across Stages of the Policy Process



* blue line represents value content in 2013 and orange line denotes value content in 2017.

Repeated T-Tests Measures

The repeated *t*-test for the time variable tests whether the frequency of the values expressed differ significantly between 2013 and 2017. As shown in Table 5.4, the t-tests for efficiency and customer service satisfaction show that the means are not the same in 2013 and 2017. Specifically, the gap between archival-based and perceptual measures between both efficiency and customer service satisfaction as represented in the agency reports and the managers' perceptions of efficiency in their agencies differed significantly between 2013 and 2017 (t-statistic=2.13 and 2.77, respectively at $p<0.05$). The gap between archival-based and perceptual measures of efficiency was less in 2017 (mean=3.18) than in 2013 (mean=3.26), and the gap between archival-based and perceptual measures of customer service satisfaction was less in 2017 (mean=3.20) than in 2013 (mean=3.30), supporting H3a and H3c.[13] This means the gap between archival-based and perceptual measures of efficiency and customer service satisfaction exist, and the size of the gaps are decreasing in 2017, compared to the size of the gaps in 2013. The changes in the gap between archival-based and perceptual measures of service quality, on the other hand, did not demonstrate significant effects over time, meaning there is no change in the size of the gap between archival-based and perceptual measures of service quality over the years. Both objective and subjective goal proximity and social proximity differed significantly between 2013 and 2017 (t-statistics=-2.98, 2.89, and -4.45, respectively at $p<0.01$). The objective goal proximity increased in 2017 (mean=0.06) compared with 2013 (mean=-0.05), while the subjective goal proximity decreased in 2017 (mean=4.31) compared with 2013

[13] H3a predicts that "the gap between archival-based and perceptual measures of efficiency will be less in 2017 than in 2013" and H3c states that "the gap between archival-based and perceptual measures of customer service satisfaction will be less in 2017 than in 2013."

(mean=4.40). The informational proximity and psychological proximity did not change significantly over time.

This study has confirmed that the gaps between archival-based and perceptual measures of efficiency and customer service satisfaction decrease as performance measurement validity increases. Based on the radar chart and repeated T-tests, even though there is value coherence between the Obama administration in 2013 and the Trump administration in 2017 and few differences between the gap levels in 2013 and 2017, the gap between archival-based and perceptual measures of value content in performance measurement decreased overall. This implies that the federal managers' experience may increase their ability to render accurate assessments of value content over time.

Table 5.4 T-Tests for Time Effects

Variable	Mean (2013)	Mean (2017)	df	t-statistic
Gap of efficiency validity	3.2603	3.1802	4286	2.1335*
Gap of service quality validity	1.8230	1.8907	4282	-1.4533
Gap of customer service satisfaction validity	3.3028	3.1973	4282	2.7721**
Objective goal proximity	-0.0529	0.0630	2651	-2.9762**
Subjective goal proximity	4.3994	4.3069	4712	2.8909**
Informational proximity	0.0449	-0.0534	2651	2.5250
Psychological proximity	0.0052	-0.0062	2651	0.2915
Social proximity	-0.0789	0.0940	2651	-4.4522***
Manager's accountability	0.0132	-0.0142	4349	0.9028
Stakeholder attention	-0.0297	0.0318	4349	-2.0280
Agency size	10.1278	10.1847	4786	-1.4016
Institutional type	0.6716	0.6761	4786	-0.3251
SES	0.1997	0.1923	4786	0.6431

*p<0.05, **p<0.01, ***p<0.001

Findings: Proximity and Performance Measurement Validity

In order to test the hypotheses formulated in Chapter 4, this study used data from FMOPMI and PARs in 2013 and 2017 by employing multi-level modeling (ML). The unit of analysis was individual managers in federal agencies. Using the available data, four dependent variables were developed in Chapter 4: performance measurement validity of efficiency, service quality, customer service satisfaction, and social equity.

Since the managers were nested within their agency, ML was employed. The dependent variables were measured by the gap between archival-based and perceptual measures of value content—e.g., efficiency, service quality, customer service satisfaction, and social equity. A lower gap between archival-based and perceptual measures reflects greater agreement and coherence between archival-based and perceptual measures. To improve model specification, fixed effects and random effects were used. Tables 5.5 through 5.8 present three different models to help determine the most accurate and improved model. Model 1 shows the fixed effects while Model 2 shows the random intercept model that implies the 23 agencies have separate intercepts. Model 3 presents the random intercept and slope model based upon whether the 23 agencies have discrete intercepts and slopes according to agency size.

Many similarities were found among the models regarding the direction and statistical significance of the main variables, providing added confidence in the model specifications and results. To find the best fitting model, the Likelihood ratio test was conducted, and AIC and BIC values were considered for comparing nested linear models. Based on the ICC value, performance measurement validity was not fully captured by Model 1, the fixed-effects model, which assumes that the same intercepts and slopes characterize all individuals in the analysis.

One way to model the different gaps in performance measurement validity was to allow each of the 23 agencies to have its own random intercept. This option seemed appropriate for this organizational study. Thus, Model 2 was added so the random intercept could allow for the possibility that the mean percent of each value varied among the agencies. While Model 1, the fixed effects model, has no specific predictors included to partition the variability in the managers' perception of performance measurement validity of the value content between-group variability, Model 2 with the random intercept includes group variability. Comparing Models 1 and 2, Tables 5.5 5.8 show an estimated standard deviation of the random intercepts (Efficiency = 0.06, Customer service satisfaction = 0.05, and Social Equity = 0.14),[14] along with a standard error (Efficiency = 0.02, Customer service satisfaction = 0.02, Social Equity = 0.06) and 95% confidence interval for that standard deviation. Based on the results, all value content's standard deviation of u_0 is significantly different from zero; Wald $\chi 2$ statistic, which explains the model fit, was increased; and AIC and BIC values were smaller in Model 2 than those in Model 1. Thus, the random intercept model (Model 2) is preferred because it provides the best fit for efficiency, customer service satisfaction, and social equity.

Model 3, the random intercept and slope model, added the organizational effect of agency size for model improvement over the previous model. Another agency-level factor, institutional type, failed to improve the model. Service quality with the organizational effect of agency size provided the best fit (LR chi2=2.38, *p*=0.3044). Based on the results, all value content's standard deviation of u_0 was different from zero, suggesting significant agency-to agency variation in the slope coefficients.

[14] The results for service quality are described in the next paragraph since service quality is analyzed in Model 3, which is a different model than that used for efficiency, customer service satisfaction, and social equity.

In summary, this study uses two models. According to the Likelihood ratio test, Model 2, the random intercept model, offered the best fit for efficiency, customer service satisfaction, and social equity (LR chi2=94.74, 66.44, and 45.87, respectively at $p<0.001$); while Model 3, the random intercept and slope model, offered the best fit for service quality (LR chi2= 497.77, $p<0.001$).

Hypothesis H4 predicted that objective proximity would have a significant influence on performance measurement validity. The results of Model 1, presented in Tables 5.5 through 5.8, show that objective goal proximity had a negative impact on the gap between archival-based and perceptual measures of social equity (β = -0.25, $p<0.001$), supporting H4d. Managers' direct involvement in goal setting decreases bias. Managers' involvement in setting performance measures and goals develops their professional skills, including an accurate understanding and articulation of the agency's mission and performance indicators. H5 predicted that the four dimensions of subjective proximity would have a negative impact on performance measurement validity. The results show that subjective goal proximity in Model 1 had a positive impact on the gap between archival-based measures and managers' perceptions for all dimensions of value content, supporting hypotheses H5a-1 through H5a-4. Especially, subjective goal proximity increased the gaps on efficiency (β = 0.18, $p<0.001$), service quality (β = 0.21, $p<0.001$), customer service satisfaction (β = 0.12, $p<0.001$), and social equity (β = 0.40, $p<0.001$). Since familiarity is a ''non-specific feeling of remembrance in the absence of conscious recollection" (Fazendeiro, Huber, Curran, and Winkielmann, 2006, p. 3), a managers' goal familiarity may lead to significant deviations from the standards of behavior predicted by the standard, increasing the individual's bias (Seiler et al., 2008).

Informational proximity was positively associated with the gap between archival-based and perceptual measures for all types of value content, supporting H5b-1, H5b-2, H5b-3, and H5b-4. Informational proximity had a consistently positive impact on the gap between archival-based and perceptual measures of efficiency ($\beta = 0.52$, $p<0.001$), service quality ($\beta = 0.43$, $p<0.001$), customer service satisfaction ($\beta = 0.49$, $p<0.001$), and social equity ($\beta = 0.43$, $p<0.001$). Neutral performance information may improve the attainment of a precisely defined organizational mission (Baloglu, 2001). Information, however, is not always neutral; it has to be produced and analyzed so that it can support certain groups or undermine others (Van de Walle & Van Dooren, 2010). Performance information does not necessarily reduce uncertainty (Learmonth & Harding, 2006) and more information may do little to improve our understanding (Bawden & Robinson, 2009; Ditton, Farrall, Bannister, & Gilchrist, 2000; Schneider, 1987). Thus, informational proximity may affect the managers' bias, increasing the gap between objective and subjective performance measures.

Psychological proximity was positively associated with the gap between archival-based and perceptual measures of all dimensions of value content, supporting H5c-1, H5c-2, H5c-3, and H5c-4. Psychological proximity was positively associated with the gap between archival-based and perceptual measures of efficiency ($\beta = 0.23$, $p<0.001$), service quality ($\beta = 0.15$, $p<0.001$), customer service satisfaction ($\beta = 0.21$, $p<0.001$), and social equity ($\beta = 0.19$, $p<0.001$). Distortion effects are also observed in managers with more perceptual similarity and closeness due to their common understandings and experience. If managers have experienced difficulties in finding meaningful performance measures and reliable data, this could undermine performance measurement validity and vice versa.

Social proximity was positively associated with the gap between archival-based and perceptual measures for all forms of value content, supporting H5d-1, H5d-2, H5d-3, and H5d-4. Particularly, social proximity was positively associated with the gap between archival-based and perceptual measures of efficiency ($\beta = 0.28$, $p<0.001$), service quality ($\beta = 0.29$, $p<0.001$), customer service satisfaction ($\beta = 0.29$, $p<0.001$), and social equity ($\beta = 0.29$, $p<0.001$). Social proximity is affected by individual or group feelings based on their interactions, and pre-existing friendships that stimulate emotional bonds (Boschma, 2005; Lundvall, 1993; Uzzi, 1997). Closer interpersonal distances affect the managers' perception not only in intergroup performance measurement but also in organizational performance measurement (Ensari & Miller, 2002). Together, the results confirm that managers with more subjective proximity are likely to render more biased assessments of value content, meaning the gap between objective and subjective measures increases.

Regarding organizational-level effects, the results for the values of efficiency, customer service satisfaction, and social equity are shown in Model 2, the random intercept model, while the results for the value of service quality is shown in Model 3, the random intercept and slope model. Organizational variables such as agency size and institutional type were included in both models. According to the Model 1, agency size was associated with the gap between archival-based and perceptual measures of value content. Institutional type did not affect the gap, but this study explored agency size as another organizational-level effect. The agency size variable was statistically significant based on the ICC value, suggesting that around 4%-27% of the variance in perceptions of value content could be explained by between-agency differences. This result confirms that agency size affects the level of measurement bias exhibited by employees when judging value content.

In addition, when holding all other variables constant, manager's accountability had no effect on the gap between archival-based and perceptual measures of efficiency, service quality, customer service satisfaction, and social equity. External stakeholder attention, such as from Congress and the audit community (e.g., the GAO and Inspectors General), was only positively related to the gap between archival-based and perceptual measures of social equity ($\beta = 0.15$, $p<0.001$); it did not affect the values of efficiency, service quality, and customer service satisfaction. External stakeholders, on the other hand, can push for the inclusion of democratic values and social responsibility to respond to competing demands from powerful politicians and knowledgeable citizens (Saltzstein, 1992). Regarding the control variables in Model 1, year was negatively associated with efficiency ($\beta = -0.11$, $p<0.001$) and customer service satisfaction ($\beta = -0.13$, $p<0.001$). This confirms that the managers' bias in value content in 2013 was larger than in 2017. Managers in the SES were not significantly different from other managers on their bias regarding performance measurement validity. Based on the results of the models, the equations of performance measurement validity were written as:

The model of $GAP_{EFF} = 0.41 - 0.05Goal_{Ob_i} + 0.18Goal_{sub_i} + 0.51Info_i + 0.23Psy_i + 0.29Soc_i + 0.05Acc_i + 0.02Att_i - 0.05SES_i - 0.10year_i - 0.06Inst_i + 0.001Size_i + u_{0j} + +\epsilon_i$

The model of $GAP_{SQ} = -1.51 + 0.01Goal_{Ob_i} + 0.19Goal_{sub_i} + 0.41Info_i + 0.17Psy_i + 0.29Soc_i + 0.04Acc_i + 0.00Att_i + 0.04SES_i + 0.06year_i + 0.26Inst_i + 0.002Size_i + u_{0j} + u_{1j}Size_{1ij} + \epsilon_i$

The model of $GAP_{CS} = 0.11 - 0.04Goal_{Ob_i} + 0.12Goal_{sub_i} + 0.48Info_{ij} + 0.21Psy_{ij} + 0.29Soc_{ij} + 0.04Acc_{ij} + 0.02Att_{ij} + 0.08SES_{ij} - 0.13year_{ij} + 0.18Inst_{ij} - 0.001Size_{ij} + u_{0j} + \epsilon_{ij}$

The model is $GAP_{SE} = 0.93 - 0.22Goal_{Ob_i} + 0.40Goal_{sub_i} + 0.42Info_{ij} + 0.21Psy_{ij} + 0.28Soc_{ij} - 0.04Acc_{ij} + 0.15Att_{ij} - 0.05SES_{ij} - 0.18Inst_{ij} + 0.002Size_{ij} + u_{0j} + \epsilon_{ij}$

To summarize, the random intercept model was applied for efficiency, customer service satisfaction, and social equity while the random intercept and slope model was employed for service quality. The results confirm that subjective proximity increases bias regardless of value content, while objective proximity decreases the gap between archival-based and perceptual measures of social equity, raising questions about why these two forms of proximity would exert different effects. This may come from the differences between actual involvement and different types of familiarity of goals, measures, and performance information. Thus, it implies that the agencies and managers must recognize the cognitive biases that influence their decision-making regarding goal setting and performance measurements. Especially, regardless of value content, informational proximity as one of the subjective proximities is the most powerful factor affecting performance measurement validity. Since performance information refers to the process of collecting and reporting performance data for effective monitoring and program management (Pollanen, 2014), performance information provides the underlying logic for decision making. At the same time, the greater the accessibility of performance information, the greater the likelihood of information abuse and misuse. Moynihan (2008b) also pointed out that performance

information is not comprehensive or objective, but always incomplete and ambiguous. Therefore, it could reduce performance measurement validity.

Table 5.5 Performance Measurement Validity of Efficiency Value

	Fixed Effects		Random Intercept		Random Intercept & Slope	
	Model 1		**Model 2**		Model 3	
	β	S.E.	β	S.E.	β	S.E.
Objective goal proximity	-0.04	0.03	-0.05	0.03	-0.05	0.03
Subjective goal proximity	0.18***	0.03	0.18***	0.02	0.18***	0.02
Informational proximity	0.52***	0.02	0.51***	0.02	0.51***	0.02
Psychological proximity	0.23***	0.02	0.23**	0.02	0.23***	0.02
Social proximity	0.28***	0.03	0.29***	0.02	0.29***	0.03
Manager's accountability	0.07*	0.03	0.05	0.03	0.05	0.03
Stakeholder attention	0.03	0.02	0.02	0.02	0.02	0.02
SES	-0.06	0.05	-0.05	0.05	-0.05	0.05
Year	-0.11**	0.04	-0.10*	0.04	-0.09*	0.04
Institutional type	-0.03	0.05	-0.06	0.14	-0.16	0.09
Agency size, (/1000)	0.001***	0.00	0.001	0.00	0.001	0.00
Constant	0.20	0.48	0.41	0.47	0.44	0.47
Individual-level variance	1.010 (0.03)		0.955 (0.03)		0.954 (0.04)	
Agency-level variance	-		0.059 (0.02)		0.001 (0.00)	
Deviance	6539.0656		6455.1671		6450.5178	
Wald χ^2	1036.87***		1039.67***		1050.51***	
Log likelihood	-3269.5328		-3227.5836		-3225.2589	
ICC	-		0.06		0.11	
AIC	6567.066		6485.167		6484.518	
BIC	6647.411		6571.251		6582.079	

*p<0.05, **p<0.01, ***p<0.001

Table 5.6 Performance Measurement Validity of Service Quality Value

	Fixed Effects		Random Intercept		Random Intercept & Slope	
	Model 1		Model 2		**Model 3**	
	β	S.E.	β	S.E.	β	S.E.
Objective goal proximity	-0.01	0.04	0.01	0.03	0.01	0.03
Subjective goal proximity	0.21***	0.03	0.19***	0.03	0.19***	0.03
Informational proximity	0.43***	0.03	0.41***	0.03	0.41***	0.03
Psychological proximity	0.15***	0.03	0.17***	0.02	0.17***	0.02
Social proximity	0.29***	0.03	0.29***	0.03	0.29***	0.03
Manager's accountability	0.04	0.04	0.04	0.03	0.04	0.03
Stakeholder attention	-0.00	0.03	0.00	0.03	0.00	0.03
SES	-0.04	0.07	0.04	0.06	0.04	0.06
Year	0.06	0.06	0.07	0.05	0.06	0.05
Institutional type	0.19*	0.07	0.32***	0.32	0.26	0.32
Agency size, (/1000)	0.001*	0.02	0.002*	0.08	0.002**	0.08
Constant	-1.31*	0.66	-1.47*	0.91	-1.51*	0.91
Individual-level variance	1.708 (0.05)		1.327 (0.04)		1.328 (0.04)	
Agency-level variance	-		0.422 (0.13)		0.425 (0.13)	
Deviance	7719.1056		7224.7116		7213.1279	
Wald χ^2	468.65***		582.56***		587.66***	
Log likelihood	-3859.5528		-3612.3558		-3606.5639	
ICC	-		0.24		0.24	
AIC	7747.106		7254.712		7247.128	
BIC	7827.414		7340.756		7344.645	

*p<0.05, **p<0.01, ***p<0.001

Table 5.7 Performance Measurement Validity of Customer Service Satisfaction Value

	Fixed Effects		Random Intercept		Random Intercept & Slope	
	Model 1		**Model 2**		Model 3	
	β	S.E.	β	S.E.	β	S.E.
Objective goal proximity	-0.06	0.03	-0.04	0.03	-0.04	0.03
Subjective goal proximity	0.12***	0.02	0.12***	0.02	0.12***	0.02
Informational proximity	0.49***	0.02	0.48***	0.02	0.48***	0.02
Psychological proximity	0.21***	0.02	0.21***	0.02	0.21***	0.02
Social proximity	0.29***	0.03	0.29***	0.03	0.29***	0.03
Manager's accountability	0.03	0.03	0.04	0.03	0.04	0.03
Stakeholder attention	0.01	0.02	0.02	0.02	0.02	0.02
SES	0.12*	0.05	0.08	0.05	0.08	0.05
Year	-0.13**	0.04	-0.13**	0.04	-0.12**	0.04
Institutional type	0.17***	0.05	0.18	0.13	0.17	0.12
Agency size, (/1000)	-0.001**	0.02	-0.001	0.04	-0.001	0.04
Constant	0.31	0.49	0.11	0.49	0.10	0.63
Individual-level variance	1.077 (0.03)		1.028 (0.03)		1.029 (0.03)	
Agency-level variance	-		0.048 (0.02)		0.031 (0.02)	
Deviance	6680.1377		6620.4369		6617.5363	
Wald χ^2	842.91***		851.59***		853.38***	
Log likelihood	-3340.0689		-3309.2047		-3308.7682	
ICC	-		0.05		0.06	
AIC	6708.138		6648.409		6651.536	
BIC	6788.483		6734.493		6749.098	

*p<0.05, **p<0.01, ***p<0.001

Table 5.8 Performance Measurement Validity of Social Equity Value

	Fixed Effects		Random Intercept		Random Intercept & Slope	
	Model 1		**Model 2**		Model 3	
	β	S.E.	β	S.E.	β	S.E.
Objective goal proximity	-0.25***	0.06	-0.22***	0.06	-0.22***	0.06
Subjective goal proximity	0.40***	0.05	0.40***	0.05	0.40***	0.05
Informational proximity	0.43***	0.05	0.42***	0.05	0.42***	0.05
Psychological proximity	0.19***	0.05	0.21***	0.04	0.21***	0.04
Social proximity	0.29***	0.06	0.28***	0.05	0.28***	0.05
Manager's accountability	-0.07	0.06	-0.04	0.06	-0.05	0.05
Stakeholder attention	0.15***	0.05	0.15***	0.05	0.14***	0.05
SES	0.05	0.11	-0.05	0.10	-0.06	0.10
Institutional type	-0.16	0.11	-0.18	0.24	-0.15	0.20
Agency size, (/1000)	0.002***	0.00	0.002**	0.00	0.003**	0.00
Constant	0.96	1.13	0.93	1.13	0.87	1.13
Individual-level variance	2.347 (0.09)		2.215 (0.09)		2.213 (0.09)	
Agency-level variance	-		0.144 (0.06)		0.05 (0.05)	
Deviance	4588.0409		4550.2884		4548.273	
Wald χ^2	215.53***		227.88***		233.47***	
Log likelihood	-2292.0179		-2275.1442		-2274.1365	
ICC	-		0.06		0.09	
AIC	4614.041		4578.288		4580.273	
BIC	4680.67		4650.042		4662.278	

*p<0.05, **p<0.01, ***p<0.001

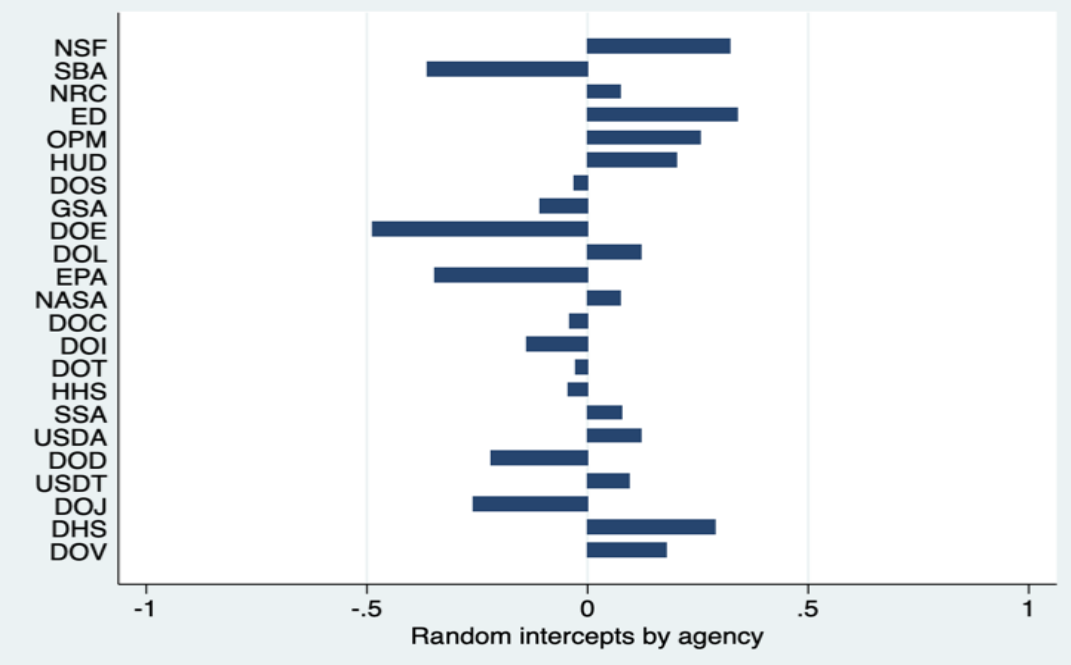
To clarify the relationship between proximity and the performance measurement validity of value content, random intercepts on efficiency, customer service satisfaction, and social equity for each agency were calculated, whereas random slopes on the service quality value for each agency were computed according to the agency size. Each agency's random intercept is shown in

a bar chart in Figures 5.2, 5.3, and 5.4. In these Figures, agencies were listed from smallest agencies to largest agencies. These intercepts do vary from place to place. Even though there are small agency's variations, these variations represent differences in bias. Interestingly, managers tend to under-estimate organizational goals' reflected efficiency and social equity value in the smallest and largest agencies, while managers tend to over-estimate organizational goals' impact on customer service satisfaction values in the smallest and largest agencies.

In addition, considering agency effects produced the best linear unbiased predictions (BLUPS) of random effects, the random slope coefficients for service quality ranged from -0.007 to 0.005, as shown in Table 5.9.[15] Since there are comparably huge agency variations in the service quality value, which is 24%, managers need to consider the agency's capacity for setting true performance measures on service quality. Organizational differences such as agency size are likely to affect the ability of managers to accurately interpret and compare the value content of performance indicators across organizational settings. This study measured agency size as the number of employees, and the results show agency size affects service quality. According to mean of the slope coefficient as shown in Table 5.9, the gap between the archival and perceived measures of service quality in the smallest and largest agencies increased as agency size increased.

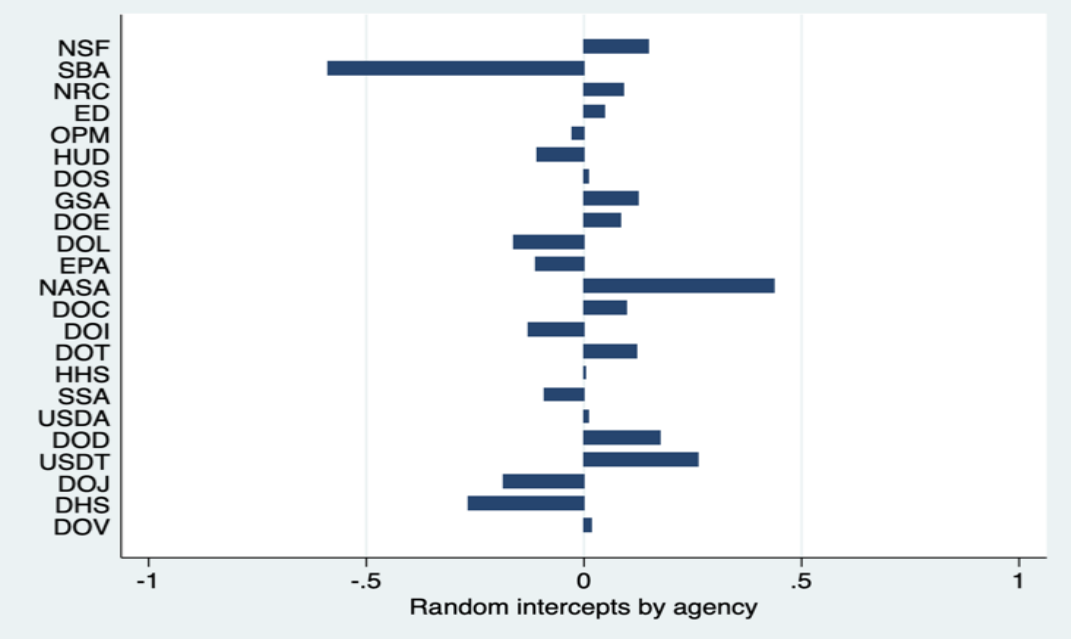
[15] To specify random coefficients on agency-level variables, a comparison test of the full model was conducted. Likelihood-ratio tests established that random coefficients on only agency size have statistically significant variation.

Figure 5.2 Random Intercept for Efficiency by Agency



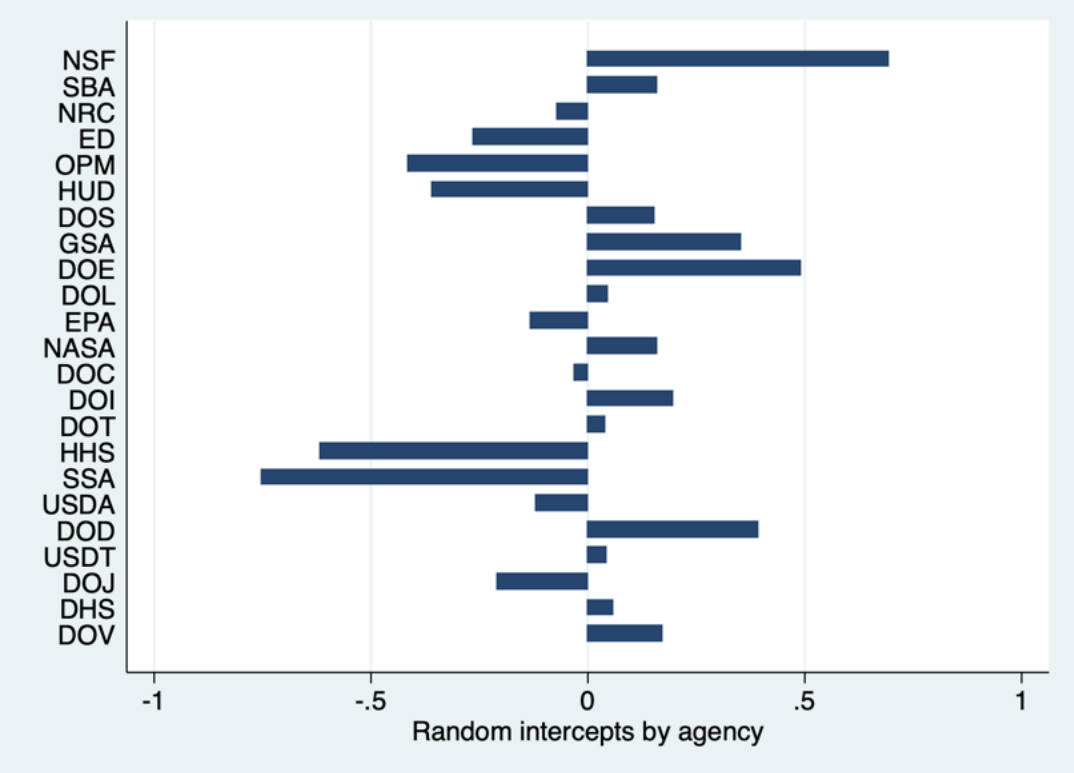
Note: Agencies are listed according to their size (from smallest to largest)

Figure 5.3 Random Intercept for Customer Service Satisfaction by Agency



Note: Agencies are listed according to their size (from smallest to largest)

Figure 5.4 Random Intercept for Social Equity by Agency



Note: Agencies are listed according to their size (from smallest to largest)

Table 5.9 Random Slope Coefficients for Service Quality Value

Agency	Mean (Random Slope)
NSF	0.0008998
SBA	0.0005564
NRC	0.004913
ED	0.0012177
OPM	-0.0019909
HUD	-0.001844
DOS	-0.0022379
GSA	-0.0071231
DOE	0.001617
DOL	-0.0003085
EPA	-0.0001058
NASA	-0.0009254
DOC	-0.0011244
DOI	0.0010154
DOT	-0.0021165
HHS	0.0000936
SSA	0.0045119
USDA	-0.0006933
DOD	0.0011741
USDT	-0.0014057
DOJ	-0.0017167
DHS	0.0021599
DOV	0.004546

Note: Agencies are listed according to their size (from smallest to largest)

CHAPTER 6

CONCLUSION

Many scholars and practitioners have called for effective performance management in the public sector (e.g., Fitzgerald, Brignall, Silvestro, & Voss, 1991; Hood, 1991; Kaplan & Norton, 1992; Kettl, 2000; Sanderson, 2001). Accordingly, performance measurement is a promising area within public management and performance management, even though it presents considerable challenges (Bouckaert & Halligan, 2007; Kamensky & Morales, 2005). Since public managers may know more about their organizations than other stakeholders (Brewer & Walker, 2012, p. 125), their assessments are valuable. As mentioned in earlier chapters, managers set organizational goals in the agency's annual performance and accountability report (e.g., the agency plan) that reflect the managers' perceptions of needs. Therefore, congruence can be expected between an agency's goals as shown in the annual reports and managers' perception of those goals. On the other hand, managers' attention, devotion, support, and commitment to performance management activities in public organizations may sometimes inflate their attitudes, perceptions, and opinions related to their organizations.

Following on this theme, this dissertation strives to develop a theoretical and practical understanding of the relationship between managers' proximity to the performance measurement process and their potential bias in assessing performance-related issues. In this study, bias is examined by understanding whether the agency's annual plan and the managers' perceptions of organizational goals are congruent, which leads to high measurement accuracy and validity. Lack

of congruence implies gaps between objective and subjective measures – between measures taken from the agency's annual plans and the managers' perceptions of the goals stated in the agency's annual plan. These gaps reflect managers' potential bias and they result in poor measurement validity. Therefore, this study focuses on explaining the gap between objective and subjective measures of value content by examining both the individual- and organizational-level impact of those differences in value content among the largest U.S. federal government agencies in the years 2013 and 2017.

In this concluding chapter, key findings of this study are summarized. Then, the theoretical and practical implications of these findings are discussed, followed by some limitations of the study and some recommendations for further research.

Summary of Findings

All hypotheses were tested to examine the relationship between managers' proximity and performance measurement validity of value content in 2013 and 2017. Since an agency's annual performance and accountability report (e.g., the agency plan) reflects the managers' perceptions of the goals of the agency, there is an implied underlying congruence between an agency's goals and the managers' perceptions of those goals. Thus, a high level of performance measurement validity can achieve measurement coherence between the measures taken from the agency's annual plans and the managers' perceptions of the goals stated in the annual plans. Otherwise, a lack of coherence implies gaps, managers' bias, and poor measurement validity.

For analyzing the relationship between performance measurement validity and managers' proximity, the main independent variables included objective goal proximity, subjective goal proximity, informational proximity, psychological proximity, and social proximity, while

additional independent variables were manager's accountability and external stakeholder attention. The dependent variable was performance measurement validity of value content consisting of four values—efficiency, service quality, customer service satisfaction, and social equity. For the purposes of this study, performance measurement validity was defined as the gap between objective (archival-based) and subjective (perceptual) measures of value content. More specifically, the gap represents the difference between managers' perceptions of value content in their agencies' performance measurement activities express in manager surveys, and archival-based measures of value content developed from intensive content analyses of those agencies' annual performance and accountability reports. Additional variables included manager accountability and stakeholder attention. Manager accountability is not related to the gap between archival-based and perceptual measures, while stakeholder attention seems to increase the gap. Control variables included manager's position (e.g. Senior Executive Service or not), and agency size and institutional type at the organizational level. Among the organizational level factors, agency size is likely to increase the gap between archival-based and perceptual measures, but institutional type does not seem to matter. Table 6.1 summarizes the results of the hypotheses formulated and tested in this dissertation.

Table 6.1 Summary of Hypotheses and Results

Hypothesis	Result
H1. The value content will be more frequently expressed in the 2017 annual agency report than in the 2013 report.	
H1a. The efficiency value will be more frequently expressed in the 2017 annual agency report than in the 2013 report.	Supported
H1b. The service quality value will be more frequently expressed in the 2017 annual agency report than in the 2013 report.	Not Supported
H1c. The customer service satisfaction value will be more frequently expressed in the 2017 annual agency report than in the 2013 report.	Supported
H1d. The social equity value will be more frequently expressed in the 2017 annual agency report than in the 2013 report.	Not Supported
H2. The value content will be more frequently expressed in the outcome stage than in the input and output stages in the annual agency report.	
H2a. The efficiency value will be more frequently expressed in the outcome stage than in the input and output stages in the annual agency report.	Supported
H2b. The service quality value will be more frequently expressed in the outcome stage than in the input and output stages in the annual agency report.	Supported
H2c. The customer service satisfaction value will be more frequently expressed in the outcome stage than in the input and output stages in the annual agency report.	Not Supported
H2d. The social equity value will be more frequently expressed in the outcome stage than in the input and output stages in the annual agency report.	Partially Supported
H3. The gap between archival-based and perceptual measures of value content will be less in 2017 than in 2013.	
H3a. The gap between archival-based and perceptual measures of efficiency will be less in 2017 than in 2013.	Supported
H3b. The gap between archival-based and perceptual measures of service quality will be less in 2017 than in 2013.	Not Supported
H3c. The gap between archival-based and perceptual measures of customer service satisfaction will be less in 2017 than in 2013.	Supported
H4. Managers' objective proximity to the performance management process will decrease the gap between archival-based and perceptual measures of value content.	
H4a. Objective goal proximity will decrease the gap between archival-based and perceptual measures of efficiency.	Not Supported
H4b. Objective goal proximity will decrease the gap between archival-based and perceptual measures of service quality.	Not Supported
H4c. Objective goal proximity will decrease the gap between archival-based and perceptual measures of customer service satisfaction.	Not Supported

H4d. Objective goal proximity will decrease the gap between archival-based and perceptual measures of social equity.	Supported
H5. Managers' diverse forms of subjective proximity to the performance management process will increase the gap between archival-based and perceptual measures of value content.	
H5a. Subjective goal proximity will increase the gap between archival-based and perceptual measures of value content.	
H5a-1. Subjective goal proximity will increase the gap between archival-based and perceptual measures of efficiency.	Supported
H5a-2. Subjective goal proximity will increase the gap between archival-based and perceptual measures of service quality.	Supported
H5a-3. Subjective goal proximity will increase the gap between archival-based and perceptual measures of customer service satisfaction.	Supported
H5a-4. Subjective goal proximity will increase the gap between archival-based and perceptual measures of social equity.	Supported
H5b. Informational proximity will increase the gap between archival-based and perceptual measures of value content.	
H5b-1. Informational proximity will increase the gap between archival-based and perceptual measures of efficiency.	Supported
H5b-2. Informational proximity will increase the gap between archival-based and perceptual measures of service quality.	Supported
H5b-3. Informational proximity will increase the gap between archival-based and perceptual measures of customer service satisfaction.	Supported
H5b-4. Informational proximity will increase the gap between archival-based and perceptual measures of social equity.	Supported
H5c. Psychological proximity will increase the gap between archival-based and perceptual measures of value content.	
H5c-1. Psychological proximity will increase the gap between archival-based and perceptual measures of efficiency.	Supported
H5c-2. Psychological proximity will increase the gap between archival-based and perceptual measures of service quality.	Supported
H5c-3. Psychological proximity will increase the gap between archival-based and perceptual measures of customer service satisfaction.	Supported
H5c-4. Psychological proximity will increase the gap between archival-based and perceptual measures of social equity.	Supported
H5d. Social proximity will increase the gap between archival-based and perceptual measures of value content.	
H5d-1. Social proximity will increase the gap between archival-based and perceptual measures of efficiency.	Supported

H5d-2. Social proximity will increase the gap between archival-based and perceptual measures of service quality.	Supported
H5d-3. Social proximity will increase the gap between archival-based and perceptual measures of customer service satisfaction.	Supported
H5d-4. Social proximity will increase the gap between archival-based and perceptual measures of social equity.	Supported

Based on NPM principles and related administrative reforms, performance measurement validity is an issue of increasing importance. This study has systematically analyzed performance measurement validity between subjective and objective measures of value content. Performance measurement validity means that performance indicators accurately measure what they are intended to measure, which is a significant issue in managing performance in the public sector. The performance measures of interest in this study are those measuring competing values (e.g., efficiency, service quality, customer service satisfaction, and social equity) based upon the assumption that values should be balanced appropriately as suggested in the Competing Values Framework and other research on public values (Ćwiklicki, 2016; Brewer, 2013; Talbot, 2006). Hence, the distribution of value content across U.S. federal government agencies was explored in this study. Furthermore, it is also important to know whether managers have accurate perceptions of value content in their agency's performance management activities, or whether they are biased and tend to underestimate or overestimate the value content. If so, in what direction and for what reason? If there are differences between the managers' perceptions of value content and what the agency really does, we need to know more about this bias and learn how to overcome it. Thus, the three research questions posed in Chapter 1 are answered below.

Research Question 1. Are there significant differences in performance measurement validity in value content when comparing measures from 2013 and 2017*?*

There is a growing body of evidence that public services implemented by government agencies should comport with public values (Nabatchi, 2012). Bourgon (2010) has claimed that the role of public organizations is to "achieve results of high public value in ways that advance civic or democratic principles" (p. 199-200). Thus, this study has checked the value content of performance measurement activities in a set of the largest U.S. federal government agencies: the values gauged are efficiency, service quality, customer service satisfaction, and social equity. These values are expressed in an independent source—the annual PARs for each agency.

Since performance measures include an agency's key concerns and the government's commitment to a set of values (Meyer & Allen, 1997), it is important to study performance measurement indicators across stages of the policy process to understand an agency's strategic aims, including their present, short-term, and long-term intentions. In this study, the researcher mounted a content analysis to sort latent values into four categories (efficiency, service quality, customer service satisfaction, and social equity) according to the policy stages-oriented perspective —of process, output, and outcome. As Bourgon (2010) notes, the program results for an agency at each stage are important because they are connected, from input to outcome. Hence, by identifying what value was expressed in which stage, we can observe the consistency and commitment of federal agencies to certain values across time.

According to Table 5.3 and Figure 5.1, service quality was the most frequently expressed value throughout the process, output, and outcome stages of most federal agencies, compared to other public values. An important proposition of President Trump's management agenda is that the U.S. federal government should focus on service excellence and improving the accountability

of government for the quality of public services. This may be little more than political rhetoric, but it does reflect the values that the Trump administration feels are important to enunciate. Recent studies have shown that the quality of public service delivery is a major determinant of successful government performance management, and that it improves government responsiveness (e.g., Bourgon, 2010; Sofyani et al., 2020). Considering the values found in each stage of the policy process, customer service satisfaction was more frequently expressed in the process stage; and efficiency, service quality, and social equity were more frequently expressed in the outcome stage. As Bourgon (2010) noted, collective actions and interests are more often found in the outcome stage, contributing to system-wide and societal impacts. Collective actions and interests are linked to direct attention toward organizational achievements rather than to inputs and procedures.

Research Question 2. Is there a gap between individual perceptions of performance measurement validity (taken from FMOPMI surveys) and archival-based measures (taken from PARs)? and if so, does proximity help explain this gap?

This study has confirmed a statistically significant gap between individual perceptions of value content and archival-based measures. This gap is sometimes referred to as the difference between objective and subjective measures. The mean gap between archival-based and perceptual measures for the efficiency and customer service satisfaction values was lower in 2017 than in 2013. This means the gap between archival-based measures and managers' perceptions of value content—for efficiency and customer service satisfaction—has been decreasing, which leads to an increase in performance measurement validity. On the other hand, the gap between subjective and objective measures of service quality did not change significantly

from 2013 and 2017.[16] Apparently the U.S. federal government has made an effort to improve performance measurement coherence and thus validity. Hart and Newcomer's (2018) analysis suggests that the Obama administration promoted transparency in performance measurement by implementing the Government Performance and Results Modernization Act of 2010 (GPRMA). At the federal level, the GPRMA requires departments and agencies to develop a balanced set of performance measures through establishing annual performance assessments of government programs (Dodaro 2011). Moreover, the Trump administration has emphasized evidence-based performance measures for federal management improvement based on the President's Management Agenda. All federal agencies are responsible for reporting performance metrics and showing progress in linking them to results in their budget documents. These consistent efforts and attention to performance measurement across the administration apparently had a significant effect on improving performance measurement coherence, increasing validity.

The proximity concept has not been utilized in previous research on performance measurement, performance management, or in the field of Public Administration in general. The concept is introduced in this study as a potentially important explanation for perceptual bias in performance assessments and thus as a contributor to public management knowledge and practice. Supporting this hope, the primary finding of this study is that diverse forms of proximity are important predictors of performance measurement validity, affecting the gaps between individual perceptions of value content and archival-based measures. Among diverse forms of proximity, objective proximity decreased the gaps between managers' perceptions of value content and archival-based measures, whereas subjective proximity increased the gaps causing greater incoherence between measures taken from the agency's annual reports and the

[16] The social equity value was not available in the 2017 FMOPMI survey, so it was examined using 2013 data only. Therefore, this study could not examine the differences of social equity between 2013 and 2017.

managers' perceptions of goals stated in the annual reports. Informational proximity, which was one of the subjective proximity types, was the most powerful factor affecting performance measurement validity while subjective goal and psychological proximity were the least influential factors based on their coefficient sizes.

The detailed results of this study show objective goal proximity, which consists of managers' actual goal-related activities, has no impact on the gap between archival-based and perceptual measures of efficiency, service quality, and customer service satisfaction. However, objective goal proximity decreases the gap between archival-based and perceptual measures of social equity (β = -0.22), increasing performance measurement validity. On the other hand, subjective goal proximity had a positive effect on efficiency (β = 0.18), service quality (β = 0.21), customer service satisfaction (β = 0.12), and social equity (β = 0.40). Subjective goal proximity, which describes the perception of goal closeness or nearness, may encourage motivation and action when individuals perceive less distance between themselves and the desired end (Cole, Balcetis, & Dunning, 2013). Subjective goal proximity predicts increases in the amount and intensity of actions in order to achieve goal-relevant objectives (Dollard & Miller, 1950), and it might increase personal closeness to the goal. In one study, individuals perceived that an agency performed better when a target was perceived to be closer to the individual and a task seemed clearer, rather than when the target was perceived as remote and the task appeared ambiguous (Stern et al., 2013).

Informational proximity had a positive impact on the gap between archival-based and perceptual measures of efficiency (β = 0.52), service quality (β = 0.43), customer service satisfaction (β = 0.49), and social equity (β = 0.43). Regarding informational proximity, if managers have easy access to performance information, they are more likely to have inflated

perceptions, which increase the gap of performance measures relative to value content. One of the reasons is that informational proximity assumes the availability of performance information and is not necessarily related to performance information use, which has received considerable attention in the public sector. In the performance measurement and management literature, the actual use of performance information in decision making is demanding and rare (de Lancer Julnes & Holzer, 2001). Another reason is that accurate and well-organized performance information may improve the attainment of a precisely defined organizational mission (Baloglu, 2001). Performance information could increase rather than reduce ambiguity (Van Dooren & Van de Walle, 2008). Excessive information sometimes creates overload and cognitive problems, while too little information constrains all aspects of public services (Bawden & Robinson, 2009; Schneider, 1987).

Psychological proximity was positively associated with the gap between archival-based and perceptual measures of efficiency (β = 0.23), service quality (β = 0.15), customer service satisfaction (β = 0.21), and social equity (β = 0.19). Psychological proximity increases the gap between archival-based and perceptual measures of performance measurement of value content. If managers can determine meaningful measures and obtain reliable data, they are more likely to render perceptions that increase the gap between archival-based and perceptual measures of value content. Psychological proximity toward a performance measurement issue predicts cognitive distance. Paradoxically, individuals can perceive conditions differently depending on their relative situation based on their position or responsibility (Wilson et al., 2008).

Social proximity was positively associated with the gap between archival-based and perceptual measures of efficiency (β = 0.28), service quality (β = 0.29), customer service satisfaction (β = 0.29), and social equity (β = 0.29). Managers sometimes have biases in

perceiving their relational ties within certain social networks and structures. Social proximity reflects the closeness among individuals or groups in society and it represents the community bond via communication and interaction of individuals within the group. Yet this closeness can produce bias and prejudice (Boschma, 2005, 2006; Bruneel et al., 2007; Matthews & Matlock, 2011).

One of this study's most interesting findings contradicts the findings of previous research. Most research findings in sociology, psychology, physics, business management, and geology have shown that managers who manage performance activities have more accurate perceptions, making decisions with more information and producing more positive results due to greater access to information and more professional experience (e.g., Balcetis, Cole, and Bisi, 2015; Cole et al., 2013; Davis, Gross, & Ochsner, 2011; Kivetz, Urminsky, & Zheng, 2006; Knoben & Oerlemans, 2006; Williams, Stein, & Galguera, 2014). The difference may come from the proximity's comprehensive concept (Lackey & Kaczynski, 2009; Scott, 1999). Proximity captures various dimensions of closeness and provides more wide-ranging perspectives, when compared to other similar concepts such as presence, co-presence, cohesiveness, and propinquity, which were the measures used in some previous studies. In this case, the findings show that managers who manage performance activities in large federal agencies have less accurate perceptions, possibly because they are gung-ho and positively oriented towards their agency's goals and values. While this type of commitment is valuable, it apparently compromises the manager's ability to render accurate assessments of the agency's portfolio of goals and values.

Research Question 3. What other factors affect the gap between archival-based and perceived measures of performance measurement validity?

While objective goal proximity, subjective goal proximity, and informational, psychological, and social proximity are main predictors of the gap between objective and subjective measures of performance measurement validity on value content, several additional factors, such as manager's accountability and external stakeholder attention, are likely to increase the bias. Specifically, manager's accountability was a positive predictor of the gap between archival-based and perceptual measures of efficiency in the fixed effects model. Managers held accountable for results are likely to produce an increased gap between archival-based and perceptual measures of performance measurement validity on efficiency. Public organizations still emphasize improving efficiency and economy in certain service domains with decreasing production costs (Andrews & Entwistle, 2013; Ferris & Graddy, 1986). In addition, external stakeholder attention has a positive impact on managers' perceptions of performance measurement validity of social equity. External stakeholder participation has a positive influence on performance because external participants can monitor and evaluate an agency's activity (Andrews, Boyne, and Walker, 2010; Andersen, Brewer, and Leisink, 2020; Brewer & Walker, 2010; Yang, 2009). According to the recent emphasis on the need for social equity in performance measures, external stakeholders tend to pay increasing attention to the social equity value (Riege & Lindsay, 2006) as it needs to be (Charbonneau & Riccucci, 2008).

Furthermore, since individual perceptions contribute to agency-level perceptions within the hierarchical organizational structure of public organizations, this study employed a multilevel model to examine within- and between-group variability. The results show that all 23 federal agencies have different levels of performance measurement validity of value content according

to agency size. For interpreting Figures 5.2, 5.3, and 5.4, the closer the bars are to the axis, the smaller the gap between managers' perceptions and archival-based measures. In the random intercept models, Figure 5.2 reveals that, at any given level of agency effects such as institutional type and agency size, DOE, ED, and NSF have a large gap between the managers' perceptions of efficiency and efficiency expressed in annual reports. Since the gap was calculated by estimating measurement validity in the agency report and subtracting managers' perceptions, a negative gap is interpreted as managers' over-estimation of value—an average of about 4.8 points higher in DOE which over-estimated efficiency measurement validity, or more than about 3.2 points lower in NSF, which under-estimated efficiency measurement validity.[17]

Regarding customer satisfaction and social equity, all 23 federal agencies have different intercepts. Again, the gap was calculated as the difference in value content of performance measures between the agency's annual report and managers' perceptions. If the gap is negative, this can be interpreted as managers' overestimation of the value content in their agency. On the other hand, if the gap is positive in the graph, this can be interpreted as managers' underestimation. According to Figure 5.3, SBA and NASA had the largest bias regarding customer service satisfaction. Managers in the SSA were likely to overestimate the customer service satisfaction value, inflating the estimates. Managers in NSF were likely to underestimate when compared to the archival-based measure. Based on Figure 5.4, managers in the SSA overestimated social equity, whereas managers in NSF underestimated it, compared to archival-based measures. [18] Another interesting finding from these figures is that managers over-estimated

[17] Since the bar is right-handed, this can be explained as underestimation of validity since the gap was calculated by estimating measurement validity in the agency report and subtracting managers' perceptions.
[18] Figure 5.3 shows that, at any given level of agency effects, the percentage of the gap between archival-based and perceived measures of customer service satisfaction averaged about 6 points higher in SBA, and more than 4.2 points lower in NASA, when compared to HHS. Figure 5.4 shows that the percentage of the gap between archival-based and perceived measures of social equity averaged about 0.7-point higher in SSA, and more than 0.6-point lower in NSF, at any given level of agency effects, compared with the median agency score.

the measurement validity of value content when it was highly related to their agency's mission, and managers under-estimated the measurement validity of value content when it was weakly related to their agency's mission. Simply, managers are more biased about values that are central to their agency's mission and activities. This is another thread of evidence that confirms the presence of bias when managers have a stake in the issue.

Regarding service quality, the agency-level variance was high and showed a better fit in the random intercept and slope model based on the likelihood ratio test, allowing each agency line to have a different slope and different effects of explanatory variables for the agency. According to the results shown in Table 5.9, although the differences in agency slopes are not enormous, this study documents that agency effects were a statistically significant predictor of the gap between archival-based and perceptual measures of service quality. Even though service quality has recently become more important to public organizations,[19] efforts to improve service quality may vary by agency as shown by the differential effects. As several authors have noted, service quality becomes more prominent with increasing calls for administrative accountability in certain agencies (Kelly, 2005; Wang, 2002).

Contributions of the Study

First, this research provides an intriguing extension of theory on performance measurement validity. Some previous investigations have focused on perceptual perspectives which are generally indistinguishable in their performance expectations. Most notable are the

[19] The evolution of performance measurement has changed "from easy-to-measure output measures to more meaningful measures of impact" (Melkers & Willoughby, 2005, p.184). The service quality value may be the key to success in measuring the meaningful impact of performance (Ghotbabadi & Baharun, 2015; Tzeng & Chang, 2011). Since the service quality value includes diverse perspectives such as reliability and responsiveness, as described in the previous chapters, it may increase public trust by providing public service professionally.

positive perceptions of managers who typically make unrealistically high predictions about their agency's performance. Since little research has focused on the accuracy of managers' perceptions for measurement validity (Starbuck & Mezias, 1996), testing subjective (perceptual) and objective (archival based) performances in the same public management study is crucial to draw reliable inferences about measurement validity. Therefore, this study analyzes performance measurement validity by estimating the gap between managers' perceptions of value content in their agencies' performance measures and independent measures taken from agency annual PARs. This analysis confirms that significant gaps exist between managers' perceptions of the values reflected in performance measurement criteria and independent assessments based on the agency's annual performance and accountability reports. These gaps are systematic, quantified measures of manager bias, which enable further empirical analysis.

Moore (2012) mentioned that performance measures require attention to philosophical, technical, managerial, and political dimensions. Regarding the managerial and technical dimensions, managers should concentrate on improving performance measures' accuracy in important dimensions of value. Furthermore, measures should "accurately capture the degree to which valued effects are occurring" (p. 13). Therefore, managers need to develop technical skills for determining whether and to what degree a particular public value is occurring in the real world of public management practice. Considering the technical and managerial dimensions, managers, especially senior managers, could play a significant role in the development of practices and standards for effective program management. Senior managers such as Program Management Improvement Officers (PMIO), Chief Operating Officers (COO), Deputy Directors for Management, Chief Performance Officers, and Chief Human Capital Officers (CHCO) are mainly responsible for developing performance management strategies through Annual Program

Portfolio Reviews, Agency Strategic Plans, and Annual Performance Plans. This action would be in line with the GPRA Modernization Act of 2010 and the Program Management Improvement Accountability Act (PMIAA) of 2016, which focused on strengthening managerial roles in managing performance through the development of practices and standards.

Performance measurement accuracy could be achieved through the development of administrative guidelines and management controls. Even though a set of standards for performance management and administrative guidelines provided by the Office of Management and Budget (OMB) has placed greater attention on federal agencies' performance measurement efforts, inconsistency was found in some aspects of the agency reports, such as objective goals. Therefore, agencies should follow the federal performance framework provided by OMB (e.g., section 200) more diligently (Mulvaney, 2018; OMB, 2017).

Second, there is scant research focusing on performance measures and value content based in the stages of the policy process or the public performance model (i.e., input-process-output-outcome). This study tracks the value content of performance measurement criteria between 2013 and 2017 and compares it across agencies. The value content of performance measurement criteria is further studied in various stages of the policy process, which reflects the performance management process (i.e., the process, output and outcome stages). Using the policy process model provides an explicit stage-based model that can be used to disaggregate performance values into several sub-components that occur over time (Boyne, 2002). This stage-based approach enables us to track each performance criterion and latent public value across the entire process. This fine-grained breakdown provides insights on the consistency of value commitment, and it enables agencies to make adjustments while moving towards their goals.

These stages cover the entire service production process, they generally unfold in sequential order, and each stage of the progression is infused with value-based judgements.

Since previous research has suggested that future research should incorporate a balanced mix of performance indicators (e.g., Charbonneau & Riccucci, 2008; Norman-Major, 2011), it is helpful to uncover some patterns across time on performance measurement validity and the values expressed in each stage. An important purpose of tracking changes in value content across time is to provide feedback for managers in an agency who make value allocation decisions in performance measurement. The value content patterns across time were shown in the Table 5.3 and Figure 5.1. Comparing the values in each stage of the performance measurement process, this study has found evidence of value coherence between 2013 and 2017. Federal agencies apparently made an effort to achieve measurement coherence across time based on strategically structured programs and activities through annual program portfolio reviews and annual performance reports under the OMB's coordination (Mulvaney, 2018). Even though the administration changed between 2013 and 2017, from Presidents Obama to Trump, the effort to achieve measurement coherence and value congruence in performance measurement continued somewhat seamlessly. This is not to say that the respective presidential administrations embraced similar values or even led the government alike, but rather that the performance management system provided continuity for federal agencies who operate across successive presidential administrations.

This result also follows from Moore's (2012) argument. He emphasized that value coherence in performance measurement is an important feature of accountability, and this effort has consistently guided government organizations towards a conception of public value in a performance measurement system. Thus, managers should concentrate on improving the

coherence of performance measures and value content as well as to improve performance measurement accuracy in important value dimensions (Moore, 2012).

Lastly, since this dissertation has sought to address the relationship between managers' perceptions of performance measurement validity regarding the four values of efficiency, service quality, customer service satisfaction, and social equity. Managers' various forms of proximity were explored, and the concept of proximity was confirmed to be an important variable affecting managers' perceptions of performance measurement criteria and thus performance measurement validity. Most previous studies in sociology, psychology, physics, business management, and geology have examined objective proximity (e.g., as being close to a goal or place, as measured in miles) and have confirmed that objective proximity improves performance (Edwards et al., 2007; Witt et al., 2008) because experiencing tasks or perceiving a target as nearer can facilitate target-related behaviors (Balcetis, Cole, & Bisi, 2015). By expanding the focus to subjective proximity, this study broadens our theoretical understanding of the proximity concept in the public management, and provides evidence of its increased predictive power. This study confirms that subjective proximity increases the gap between archival-based and perceptual measures regardless of the value content, always reducing performance measurement validity. Objective proximity decreases the gap between archival-based and perceptual measures of social equity, increasing the gap. We can conclude that subjective and objective proximity produce different and sometimes opposite effects.[20] Managers' perception and familiarity with agency missions, and their involvement in performance-related activities regarding the certain values, may affect their perceptions of performance criteria and alter performance measurement validity.

[20] Even though objective proximity is statistically significant on only social equity, it tends toward a negative relationship on other values.

Further attention is still needed to understand the different directions of proximity effects on performance measurement validity since public managers are mainly involved in performance measurement and are tasked with managing their organizations for value creation (Brewer 2013; Andersen et al, 2020). Therefore, recognizing how proximity affects performance measurement validity is helpful for researchers and practitioners who are trying to develop more valid, reliable, and complete measures of performance in the public sector. This work illustrates alternative ways to study performance measurement validity and value content in public organizations and provides insights that will help strengthen performance management systems in government.

Practical Implications

This dissertation has implications for government officials who oversee performance management activities. Given the current allocation of value content in performance measurement in the U.S. federal government, the gap between archival-based and perceptual measures of efficiency and customer service satisfaction decreased from 2013 and 2017; on the other hand, performance measurement validity on efficiency and customer service satisfaction measures increased over time. Yet this study confirms that a persistent gap exists between archival-based and perceptual measures, and that smaller gaps exist for service quality but there are larger variations across agencies. Estimating the gap between archival-based and perceptual measures provides actionable information for policymakers and public managers overseeing the performance management process. It could, for example, enable them to see these disparities and take action to both improve agency value allocation and partially correct or adjust the distorted instrumentation that plagues perceptual measures, which are otherwise easier to collect and more

encompassing than most purported objective measures (Simon, 1956; Brewer 2007). Despite these virtues, perceptual measures are somewhat distorted.

In addition, group differences were found among federal agencies in the relationship between proximity and the gap between archival-based and perceptual measures of value content, especially on service quality and social equity in the multilevel model. This finding provides strong evidence that a gap between archival-based and perceptual measures that represents a discrepancy in measurement validity is not driven solely by an individual's proximity but is determined by both individual (i.e., proximity) and agency (i.e., agency size) effects. Nonetheless, proximity, as a significant predictor of the gap between objective and subjective measures, provides useful evidence on the degree to which federal agency performance measures are distorted in the eyes of federal agency managers due to their proximity or distance from the performance management process.

Limitations and Future Research Directions

Several limitations of this study should be noted for future research. First, previous studies in sociology, psychology, physics, business management, and geology have confirmed that proximity is a strong predictor of performance (Oerlemans & Meeus, 2005; Shepperd, Ouellette, & Fernandez, 1996). Since the research model in this study is mainly designed to examine diverse types of proximity which may suffer from ambiguity (Boschma, 2006), the findings of this study cannot be generalized to all proximity research. One source of ambiguity is that different labels are used for the same type of proximity in various disciplines. For example, social proximity is sometimes denoted as personal proximity (Schamp et al. 2004), relational proximity (Coenen et al. 2004), or part of organizational proximity (Filippi & Torre 2003).

Geographical proximity is sometimes denoted as spatial proximity (Oerlemans & Meeus, 2005) or territorial proximity (Salazar & Marin, 1977). Another source of ambiguity is that many scholars warn that different forms of proximity are correlated and overlapping, especially the relationships between cultural and institutional proximity, and organizational and institutional proximity. For example, institutional proximity portrays the extent of shared norms, habits, and rules between agents and cultural proximity describes a pattern of thoughts, feelings, and behaviors that lead to a similar interpretation of the situation (Knoben & Oerlemans, 2006). Hofstede (2001) points out that since cultures strongly influence institutions, the two concepts – cultures and institutions – are co-related. Some scholars even consider institutions to be 'cultural artifacts' (e.g., Morgan, 1997, p. 493). While this study did not include cultural, institutional, or organizational proximity to avoid those ambiguities, further study should more precisely define each type of proximity and use accurate measures of each relevant characteristic. This could lead to a crisper operational definition of proximity and a better understanding of its cascading effects. Future research should also continue to focus on the relationship between proximity and performance in different settings, such as at different levels of government, in different service sectors, and during different time periods.

Second, it is difficult to design questionnaires to obtain good objective or subjective data and low correlations between them could indicate poor measurement reliability (Starbuck & Mezias, 1996). Fortunately, this study reports a moderately high degree of agreement between manual methods and ATLAS.ti qualitative analysis software on archival-based and presumably objective performance measures. Yet there is still a concern about the similarity between subjective and objective measures. For example, the subjective dependent variable for service quality comes from the survey questionnaire, and performance measures tell us about the quality

of the products or services provided. Through content analysis, we were able to discern the values being prioritized and promoted in these activities. On the other hand, the objective dependent variable for service quality is measured by the frequencies of certain words (e.g., *qualified, high-quality, improper, error, reliable, certified, accuracy, good condition,* etc.) expressed in the annual agencies' PARs in 2013 and 2017. The information drawn from these two sources is not a perfect match, but once quantified, they should produce comparable metrics. Reflecting on these underlying measurement issues, we can infer that future research should strive for better subjective and objective measures.

Third, this study acknowledges the potential disconnect between managers' perceptions of value content and annual agency plans at some point. Some agencies may not distill their missions and value intentions into a truly accurate agency plan. In this case, managers might be more correct about their agency's values than are the agency plans. In other words, managers may know more about what their agency is actually intending and doing, rather than how those intentions and actions are portrayed in the agency plan. Yet the gap still exists. Therefore, this study does not attribute all of the gap to managers' bias. Rather, the gap means the difference between objective and subjective measures, with the proviso that some objective measures may not be fully accurate.

Besides, there is a construct validity issue in measuring dependent variables from the survey because secondary data are used. This study uses one survey questionnaire for each value content measure to estimate the managers' perceptions of performance measurement validity. Since measuring performance measurement validity is not simple due to the contested concepts of performance and measurement validity, future research needs to include additional measures of the key variables so that the results of next-generation studies will be more convincing and

refined. Another construct validity concern is that this study used aggregated mean perceptions of performance measurement validity of value content in order to compare objective measures with subjective measures across agencies. Many scholars point out that averaged or otherwise aggregated data can produce false estimates due to the non-linear nature of ordinal rankings used in survey questionnaires (Quinn, 2004; Whitford et al., 2010). Yet several studies have confirmed that averaged perceptions are fairly accurate although almost all individuals perceive inaccurately (Dawes, 1977; Starbuck & Bass, 1967). In other words, these studies suggest that means actually reconcile inaccurate responses and hew toward collective accuracy. Future studies might bypass this problem by including items that solicit continuous responses so that more reliable results can be obtained.

Fifth, in this study, the gap of performance measurement validity on efficiency, service quality, and customer service satisfaction is compared between 2013 and 2017, but the social equity value cannot be compared between 2013 and 2017 due to its omission from the latter survey. Future research needs to explore social equity in 2017 and expand the range of values to include the concepts such as responsibility, sustainability, accountability, etc. Regarding the demographic variables in the FMOPMI survey, any information that may identify respondents is redacted to ensure their anonymity (as per 4 C.F.R. § 81.6(l), which states that GAO will not release such information). Therefore, future research needs to include more diverse individual factors, such as gender, tenure, education, income level, etc. Another suggestion is that research should extend the timeline beyond two years. More data points would provide a sounder foundation for assessing change over time.

Furthermore, this study focuses on managers in government agencies because they are strategic actors in the performance management process. This study confirms that managers'

cognitive feelings and perceptions are affected by subjective proximity (e.g., Witt, Linkenauger, & Proffitt, 2012). Yet this study does not include non-managers in public organizations, and thus the findings cannot be generalized to all employees in the public sector. For example, Brewer (2005) found that supervisors tend to report more positive and optimistic responses about work-related issues and performance than street-level employees who work on the front lines (also see Brewer & Walker, 2013; Walker & Brewer, 2008). High-level political appointees were generally excluded from the surveys studied in this dissertation, and little information was available on managers' level of responsibility. Hence, further study should incorporate diverse types and levels of public employees to obtain a comprehensive understanding of proximity effects on performance measurement validity. It may be that past research using elite survey respondents has concentrated on the views of those managers who are most likely to render biased assessments on management and performance (Brewer 2006).

Lastly, given that individuals' perceptions differ from organization to organization due to cultural and institutional disparities, this study includes the agency-level effect. There is no claim that agency-level effects are the only factor influencing the gap between archival-based and perceptual measures, but it does suggest that researchers and practitioners should pay attention to not only individual-level variables but also agency-level variables in order to understand the factors contributing to an agency's performance measurement validity. This suggests that more multi-level studies will be needed, and other research strategies such as panel studies and controlled experimentation will likely provide additional insights. In addition, since there are few systematic patterns in the agency-level results from the figures and tables, and there is some evidence of an agency size effect when comparing the smallest and largest agencies, a quadratic agency size relationship could be examined in future research. Hence, future research should try

to replicate and extend the findings of this study by using a design that considers diverse agency-level variables, such as an agency's strategic planning capacity and its autonomy or independence, and also try to isolate both across-time, quadratic, and cause-and-effect relationships.

REFERENCES

Agus, A., Barker, S., & Kandampully, J. (2007). An exploratory study of service quality in the Malaysian public service sector. *International Journal of Quality & Reliability Management*, 24(2), 177-190.

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. *In Action control* (pp. 11-39), Springer, Berlin, Heidelberg.

Alach, Z. (2016). Towards a standard conceptual typology of public sector performance measurement. Kōtuitui: *New Zealand Journal of Social Sciences Online*, 12(1), 56-69.

Albouy, V., Godefroy, P., & Lollivier, S. (2010). Measuring quality of life. *Insee Références France, portrait social*.

Allen, T. J. (1977). *Managing the flow of technology: Technology transfer and the dissemination of technological information within the R&D organisation*. Cambridge, MA: The Massachusetts Institute of Technology.

Amin, A., & Wilkinson, F. (1999). Learning, proximity and industrial performance: An introduction. *Cambridge Journal of Economics*, 23(2), 121-125.

Ammons, D. N. (1995). Overcoming the inadequacies of performance measurement in local government: The case of libraries and leisure services. *Public Administration Review*, 55(1), 37-47.

Andersen, L. B., Brewer, G. A., & Leisink, P. L. M. (2020). Stakeholders, public value(s) and public service performance. In Leisink, P.L.M., Andersen, L.B., Brewer, G.A., Jacobsen,

C.B., Knies, E., & Vandenabeele, W., eds. *Managing for Public Service Performance: How People and Values Make a Difference*. Oxford University Press.

Anderson, C. R. (1988). *Management* (2nd ed): Allyn & Bacon Inc., Boston, MA.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89, 369-406.

Andrews, R. (2014). Performance management and public service improvement: Evidence review prepared for the commission on public service governance and delivery. Public Policy Institute for Wales, *PPIW Report* (3).

Andrews, R., & Boyne, G. A. (2010). Capacity, leadership, and organizational performance: Testing the black box model of public management. *Public Administration Review*, 70(3), 443-454.

Andrews, R., Boyne, G. A., & Walker, R. M. (2006). Strategy content and organizational performance: An empirical analysis. *Public Administration Review*, 66(1), 52–63.

Andrews, R., Boyne, G. A., & Walker, R. M. (2010). Administrative or survey data for measuring organizational performance: What's the difference? Paper presented at the 28 14th *International Research Society for Public Management (IRSPM) Conference*, Berne, Switzerland.

Andrews, R., Boyne, G.A., Meier, K.J., O'Toole, Jr. L. J., & Walker, R.M. (2005) Representative bureaucracy, organizational strategy, and public service performance: An empirical analysis of English local government. *Journal of Public Administration Research and Theory*, 15(4), 489-504.

Andrews, R., & Entwistle, T. (2013). Four faces of public service efficiency: What, how, when and for whom to produce. *Public Management Review*, 15(2), 246-264.

Andrews, R., & Van de Walle, S. (2013). New public management and citizens' perceptions of local service efficiency, responsiveness, equity and effectiveness. *Public Management Review*, 15(5), 762-783.

Ammons, D. N., & Rivenbark, W. C. (2008). Factors influencing the use of performance data to improve municipal services: Evidence from the North Carolina benchmarking project. *Public Administration Review*, 68(2), 304-318.

Ansoff, I. H. (1979). Strategic management. London: Macmillan.

Ansoff, I. H. (1987). The emerging paradigm of strategic behavior. *Strategic Management Journal*, 8, 501-515.

Aristigueta, M. P. (1999). *Managing for results in state government*. Greenwood Publishing Group.

Aristigueta, M. P., Cooksy, L. J., & Nelson, C. W. (2001). The role of social indicators in developing a managing for results system. *Public Performance & Management Review*, 254-269.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, 64(6), 359-372.

Babiker, I. E., Cox, J. L., & Miller, P. M. (1980). The measurement of cultural distance and its relationship to medical consultations, symptomatology and examination performance of overseas students at Edinburgh University. *Social Psychiatry*, 15(3), 109-116.

Balaboniene, I., & Vecerskiene, G. (2015). The aspects of performance measurement in public sector organization. *Social and Behavioral Sciences*, 213, 314-320.

Balcetis, E., Cole, S., & Bisi, A. (2015). Near and nothing to it: Perceived proximity improves exercise by increasing feasibility appraisals. *Motivation Science*, 1(4), 201-218.

Baloglu, S. (2001). Image variations of Turkey by familiarity index: Informational and experiential dimensions. *Tourism Management*, 22(2), 127-133.

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology*, 41(3), 586-598.

Banks, C. G., & Murphy, K. R. (1985). Toward narrowing the research-practice gap in performance appraisal. *Personnel Psychology*, 38(2), 335-345.

Bawden, D., & Robinson, L. (2009). The dark side of information: overload, anxiety and other paradoxes and pathologies. *Journal of Information Science*, *35*(2), 180-191.

Behn, R. (2003). Why measure performance? Different purposes require different measures. *Public Administration Review*, 63(5), 586-604.

Bender, K. A., & Heywood, J. S. (2006). Job satisfaction of the highly educated: The role of Gender, academic tenure, and earnings, *Scottish Journal of Political Economy*, 53(2), 253-279

Bergadaà, M., & Del Bucchia, C. (2009). La recherche de proximité par le client dans le secteur de la grande consommation alimentaire. *Management & Avenir*, 21, 121.

Berman, E. M. (1997). Dealing with cynical citizens. *Public Administration Review*, 57(2), 105-112.

Berman, E. M. (1998). *Productivity in Public and Nonprofit Organization: Strategies and Techniques*. Thousand Oaks, CA: SAGE.

Bernard, T., & Tafesse, A. S. (2014). Aspirations: An approach to measurement with validation using Ethiopian data, *Journal of African Economies*, 23(2), 189-224

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement, *Lancet*, 1, 307-310

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (p. 349–381). Jossey-Bass.

Bloom, H. S., Hill. C. J., & Riccio, J. A. (2003). Linking program implementation and effectiveness: Lessons from a pooled sample of welfare-to-work experiments, *Journal of Policy Analysis and Management,* 22(4), 551-575.

Bobak, C. A., Barr, P. J., & O'Malley, A. J. (2018). Estimation of an inter-rater intra-class correlation coefficient that overcomes common assumption violations in the assessment of health measurement scales. *BMC Medical Research Methodology*, 18(1), 93.

Bogan, C. E. & English, M. J. (1994), *Benchmarking for best practice: Winning through innovative adaptation*, New York: McGraw Hill.

Bok, D. (2001). *The trouble with government*. Cambridge, MA: Harvard University Press.

Booker, K., Gilpatric, S. M., Gronberg, T., & Jansen, D. (2007). The impact of charter school attendance on student performance. *Journal of Public Economics,* 91(5-6), 849-876.

Boschma, R. A. (2005). Proximity and innovation: A critical assessment. *Regional Studies* 39(1), 61-74.

Boschma, R. A., & Lambooy, J. G. (1999). The prospects of an adjustment policy based on collective learning in old industrial regions. *GeoJournal*, 49(4), 391-399.

Bouckaert, G. (1992). Productivity analysis in the public sector: The case of the fire service. *International Review of Administrative Sciences*, 58(2), 175-200.

Bouckaert, G., & Balk, W. (1991). Public productivity measurement: Diseases and cures. *Public Productivity & Management Review,* 15(2), 229-235.

Bouckaert, G., & Halligan, J. (2007). *Managing performance: International comparisons*. Routledge.

Bouckaert, G., & Peters, B. G. (2002). Performance measurement and management: The Achilles' heel in administrative modernization. *Public Performance & Management Review*, 25(4), 359-362.

Bourgon, J. (2010). The history and future of nation-building? Building capacity for public results, *International Review of Administrative Sciences*, 76(2), 197-218

Bovaird, T., & Löffler, E., (1996) *Public management and governance,* London and Newyork: Taylor & Francis Group, 138-144.

Bovens, M., Schillemans, T. & Hart, P. T. (2008). Does public accountability work? An assessment tool, *Public Administration,* 86(1): 225-242.

Boyne, G., & Dahya, J. (2002). Executive succession and the performance of public organizations. *Public Administration*, 80(1), 179-200.

Boyne, G. A. (2002). Concepts and Indicators of Local Authority Performance: An Evaluation of the Statutory Frameworks in England and Wales. *Public Money and Management*, 22(2), 17-24.

Boyne, G. A. (2003). Sources of public service improvement: A critical review and research agenda. *Journal of Public Administration Research and Theory*, 13(3), 367-394.

Boyne, G. A. & Walker, R. M. (2005). Introducing the "determinants of performance in public organizations" symposium. *Journal of Public Administration Research and Theory* 15(4), 483-488.

Bozeman, B., & Straussman, J. D. (1990). *Public management strategies: Guidelines for managerial effectiveness*. Jossey-Bass.

Bradshaw, M. (2011). Multiple proximities: culture and geography in the transport logistics of newsprint manufactured in Australia. *Environment and Planning* A 33, 1717-1739.

Brady, M. K., & Cronin Jr, J. J. (2001). Some new thoughts on conceptualizing perceived service quality: a hierarchical approach. *Journal of Marketing*, *65*(3), 34-49.

Breslow N. E., & Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, 88(421):9-25.

Brewer, D. D. (2011). Biases in perceiving one's own social position and social ties as evolved psychological mechanisms. Available at SSRN: http://dx.doi.org/10.2139/ssrn.1879624

Brewer, G. A. (2000). Administrative reform and organizational change in the public sector (Democracy Project AP 24/2000). Copenhagen, Denmark: University of Copenhagen.

Brewer, G. A. (2005). In the eye of the storm: Frontline supervisors and federal agency performance. *Journal of Public Administration Research and Theory* 15(4), 505-527.

Brewer, G. A. (2006). All measures of performance are subjective: More evidence on US federal agencies. In George A. Boyne, Kenneth J. Meier, Laurence J. O'Toole, Jr., and Richard M. Walker (Eds.), *Public service performance: Perspectives on measurement and management*. Cambridge, UK and New York: Cambridge University Press, pp. 35-54.

Brewer, G. A. (2013). Public management contributions for improving social service performance: Public values, public service motivation, and red tape. In Peter L. M. Leisink, Maarten van Bottenburg, Paul Boselie, and Dian Marie Hosking, (Eds.), *Managing social issues: A public values perspective: Cheltenham and Camberley,* UK: Edward Elgar Publishing Ltd., pp. 19-36.

Brewer, G. A., & Selden, S. C. (2000). Why elephants gallop: Assessing and predicting organizational performance in federal agencies. *Journal of Public Administration Research and Theory*, 10(4), 685-711.

Brewer, G. A., & Walker, R. M. (2010). The impact of red tape on governmental performance: An empirical analysis. *Journal of Public Administration Research and Theory*, 20(1), 233-257.

Brewer, G. A., & Walker, R. M. (2012). Personnel constraints in public organizations: The impact of reward and punishment on organizational performance, *Public Administration Review*, 73(1), 121-131.

Brown Jr., B. W., Lucero, R. J., & Foss, A. B. (1962). A situation where the Pearson correlation coefficient leads to erroneous assessment of reliability, *J. Clin Psychol*, 19, 95-97

Brown, M. (1996). *Keeping score: Using the right metrics to drive world-class performance*: Amacom.

Brown, R. E., & Pyers, J. B. (1988). Putting teeth into the efficiency and effectiveness of public services. *Public Administration Review*, 48(3), 735-742.

Brown, T. L., & Potoski, M. (2003). Transaction costs and institutional explanations for government service production decisions. *Journal of Public Administration Research and Theory*, 13(4), 441-468.

Brown, T. L., Potoski, M., & Van Slyke, D. M. (2006). Managing public service contracts: Aligning values, institutions, and markets. *Public Administration Review*, 66(3), 323-331.

Bruneel, J., Spithoven, A., & Maesen, A. (2007). Building trust: a matter of proximity? Paper presented at the *Babson College Entrepreneurship Research Conference*.

Bruton, A., Conway, J. H., & Holgate, S. T. (2000). Reliability: what is it, and how is it measured? *Physiotherapy*, 86-94.

Bryk, A. S., & Raudenbush, S. W. (1989). Methodology for cross-level organizational research. *Research in Organizational Behavior*, 7, 233–273.

Buchanan, B. (1975). Red tape and the service ethic: Some unexpected differences between public and private managers. *Administration and Society*, 6(4), 423-438.

Burgoon, J. K., Bonito, J. A., Ramirez Jr, A., Dunbar, N. E., Kam, K., & Fischer, J. (2002). Testing the interactivity principle: Effects of mediation, propinquity, and verbal and nonverbal modalities in interpersonal interaction. *Journal of Communication*, 52(3), 657-677.

Cairncross, F. (2002). The death of distance, *Royal Society for the Encouragement of Arts and Commerce*. 149(5502), 40-42.

Cameron, K. S. & Whetten, D. A. (1983). *Organizational effectiveness: A comparison of multiple models*. New York: Academic Press.

Capello, R. (1999). Spatial transfer of knowledge in high technology milieux: Learning versus collective learning processes. *Regional Studies*, 33(4), 353-365.

Carver, D. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control-theory approach to human behavior*. New York: Springer-Verlag.

Cassidy, J., & Kendis, S. (2011). Performance measurement: an index approach for federal agencies. Paper presented at *the PMI® Global Congress* 2011, North America, Dallas, TX.

Cha, M., Park, J. G., & Lee, J. (2014). Effects of team member psychological proximity on teamwork performance. *Team Performance Management*, 20(1/2), 81-96.

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models, *Journal of Applied Psychology*, 83(2), 2324-2246.

Chandler, A. D. (1962). *Strategy and structure: Chapters in the history of American enterprise*. Cambridge, MA: MIT Press.

Charbonneau, E. Riccucci, N. M. (2008). Beyond the usual suspects: An analysis of the performance measurement literature on social equity indicators in policing, *Public Performance & Management Review*, 31(4), 604-620.

Cheema, A., & Bagchi, R. (2011). The effect of goal visualization on goal pursuit: Implications for consumers and managers. *Journal of Marketing*, 75(2), 109-123.

Chitwood, S. R. (1974). Social equity and social service productivity. *Public Administration Review*, 34(1), 29-35.

Cho, S., & Mor Barak, M. E. (2008). Understanding of diversity and inclusion in a perceived homogeneous culture: A study of organizational commitment and job performance among Korean employees, *Administration in Social Work*, 32(4), 100-126.

Christensen, T., & Laegreid, P. (2001). (Eds.). *New public management: The transformation of ideas and practice*. Aldershot, UK: Ashgate.

Chun, Y. H., & Rainey, H. G. (2005). Goal ambiguity and organizational performance in US federal agencies. *Journal of Public Administration Research and Theory*, 15(4), 529-557.

Clay, J. A., & Bass, V. (2002). Aligning performance measurement with key management processes. *Government Finance Review*, 18(2), 26-29.

Coenen, L., Moodysson, J., & Asheim, B. T. (2004). Nodes, networks and proximities: on the knowledge dynamics of the Medicon Valley biotech cluster. *European Planning Studies*, 12(7), 1003-1018.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, S. (2008). *The responsible contract manager: Protecting the public interest in an outsourced world*: Georgetown University Press.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: a new perspective on learning an innovation. *Administrative Science Quarterly* 35(1), 128-152.

Cole, M. S., Balcetis, E., & Dunning, D. (2013). Affective signals of threat increase perceived proximity. *Psychological Science* 24(1), 34-40.

Cole, M. S., Bedeian, A. G., Hirschfeld, R. R., & Vogel, B. (2010). Dispersion-composition models in multilevel research: A data-analytic framework, *Organizational Research Methods*, 14(4), 718-734

Cramton, C. D. (2001). The mutual knowledge problem and its consequences for dispersed collaboration. *Organization Science*, 12(3), 346-371.

Curry, A., & Herbert, D. (1998). Continuous improvement in public services-a way forward. *Managing Service Quality: An International Journal*, 8(5), 339-349.

Ćwiklicki, M. (2016). Comparison of public value measurement frameworks. *Zarządzanie Publiczne*, 35(1), 20-31.

Dahl, M. S., & Pedersen, C. Ø. (2004). Knowledge flows through informal contacts in industrial clusters: myth or reality? *Research Policy*, *33*(10), 1673-1686.

Dahl, R. A. (1947). The science of public administration: Three problems. *Public Administration Review*, 7(1), 1-11.

Dahlberg, J. S. (1966). *The New York bureau of municipal research New York*: New York University Press.

Davids, M., & Frenken, K. (2018). Proximity, knowledge base and the innovation process: towards an integrated framework, *Regional Studies*, 52(1), 23-34.

Davis, J. I., Gross, J. J., & Ochsner, K. N. (2011). Psychological distance and emotional experience: What you see is what you get. *Emotion*, 11(2), 438.

Dawes, R. M. (1977). Suppose we measured height with rating scales instead of rulers, *Applied Psychological Measurement*, 1(2), 267-273

De Bruijn, H., & Dicke, W. (2006). Strategies for safeguarding public values in liberalized utility sectors, *Public Administration*, 84(3), 717-735.

De Lancer, J. P. (2008). *Performance-based management systems: Effective implementation and maintenance*. CRC Press.

De Lancer, J. P. & Holzer, M. (2001). Promoting the utilization of performance measures in public organizations: An empirical study of factors affecting adoption and implementation. *Public Administration Review*, 61(6), 693-708.

DeNisi, A. S. (2000). Performance appraisal and performance management, in Katherine, J. K., & Steve W. J. (eds). *Multilevel Theory, research, and methods in organization*s: Foundations, extensions and new directions. San Francisco: Jossey-Bass.

DeNisi, A. S., Cafferty, T. P., & Meglino, B. M. (1984). A cognitive view of the performance appraisal process: A model and research propositions. *Organizational Behavior & Human Performance*, 33(3), 360-396.

DeNisi, A. S., & Pitchard, R. D. (2006). Performance appraisal, performance management and improving individual performance: A motivational framework. *Management and Organization Review*, 2(2), 253-277.

Denhardt, J. V., & Denhardt, R. B. (2015). The new public service revisited. *Public Administration Review*, 75(5), 664-672.

Denhardt, R. B., & Denhardt, V. D. (2000). The new public service: Serving rather than steering. *Public Administration Review*, 6(6), 549-559.

Dodaro, G. L. (2011). *Government performance: GPRA Modernization Act provides opportunities to help address fiscal, performance, and management challenges: Testimony Before the committee on the budget, US Senate*. US Government Accountability Office.

Dollard, J., & Miller, N. E. (1950). *Personality and psychotherapy; an analysis in terms of learning, thinking, and culture*. McGraw-Hill Book Company, New York

Donabedian, A. (1989). Institutional and professional responsibilities in quality assurance. *International Journal for Quality in Health Care*, 1(1), 3-11.

Downey H. K., Hellriegel, D., & Slocum, J. W. Jr. (1975). Environmental uncertainty: the construct and its application, *Administrative Science Quarterly*, 2, 613-629

Downs, A. (1967). *Inside bureaucracy*: Boston: Little, Brown.

Downs, G. W., & Larkey, P. D. (1986). *The search for government efficiency*: Philadelphia: Temple University Press.

Dubnick, M. (2005). Accountability and the promise of performance: In search of the mechanisms. *Public Performance and Management Review* 28(3), 376-417.

Dubnick, M. J., & Frederickson, H. G. (2009). Accountable agents: Federal performance measurement and third-party government. *Journal of Public Administration Research and Theory*, 20, i143-i159.

Dufeu, I., & Ferrandi, J. M. (2012). Perceived proximity, rust, satisfaction and relationship commitment of members of community supported agriculture. Paper presented at the Presented at *the International Marketing Trends Conference*, Venice.

Dunleavy, P., & Hood, C. (1994). From Old Public Administration to New Public Management. *Public Money & Management*, 14(3), 9-16.

Dusenbury, P. (2000). *Governing for results and accountability: Strategic planning and performance measurement*: The Urban Institute.

Duyar, I., Gumus, S. & Sukru Bellibas, M. (2013), Multilevel analysis of teacher work attitudes: The influence of principal leadership and teacher collaboration, *International Journal of Educational Management*, 27(7), 700-719.

Edquist, C., & Johnson, B. (1996). *Institutions and organizations in systems of innovation*. Univ.

Edwards, B., Waterhouse, J., Atkinson, G., & Reilly, T. (2007). Effects of time of day and distance upon accuracy and consistency of throwing darts. *Journal of Sports Science*, 13, 1531-1536.

Eisenhardt, K. M. (1985). Control: Organizational and economic approaches. *Management Science*, 31, 134-149.

Ely, R., J. (2004). A filed study of group diversity, participation in diversity education programs, and performance, Journal of Organizational Behavior, 25(6), 755-780

Ensari, N., & Miller, N. (2002). The out-group must not be so bad after all: the effects of disclosure, typicality, and salience on intergroup bias. *Journal of Personality and Social Psychology*, 83(2), 313.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society: Series A (General)*, 120(3), 253-281.

Favero, N., Anderson, C. C., Meier, K. J., O'Toole, L. J., & Winter, S. (2015). *Is the performance effect of management underestimated? Comparing public managers' and front-line employees' perceptions of management*. Paper presented at 86th Annual Conference of the Southern Political Science Association, SPSA 2015, New Orleans, United States.

Favero, N. & Bullock, J. B. (2014). How (not) to solve the problem: An evaluation of scholarly response to common source bias, *Journal of Public Administration Research and Theory*, 25, 285-308.

Fazendeiro, T., D.E. Huber, T. C., & Winkielmann, P. (2006). Perceptual fluency effects on episodic familiarity: Decision bias versus false memory. working paper, Dept. of Psychology at University of Denver

Feldman, D. C. (2002). Managers' propensity to work longer hours: A multilevel analysis, *Human Resource Management Review*, 12(3), 339-357.

Feldman, R. (2013). Filled radar charts should not be used to compare social indicators. *Social indicators research*, 111(3), 709-712.

Fernandez, S., & Moldogaziev, T. (2011). Empowering public sector employees to improve performance: does it work? *The American Review of Public Administration*, 41(1), 23-47.

Ferris, J., & Graddy, E. (1986). Contracting out: For what? With whom? *Public Administration Review*, 332-344.

Festinger, L. (1951). Architecture and group membership. *Journal of Social Issues*, 7(1-2), 152-163.

Fiegenbaum, A., Hart, S., & Schendel, D. (1996). Strategic reference point theory. *Strategic Management Journal*, 17, 219-235.

Filippi, M., & Torre, A. (2003). Local organizations and institutions: How can geographical proximity be activated by collective projects? *International Journal of Technology Management*, 26(2-4), 386-400.

Finholt, T. A., Sproull, L., & Kiesler, S. (2002). Outsiders on the inside: Sharing know-how across space and time. *Distributed Work*, 357-380.

Finer, H. (1941). Administrative responsibility in democratic government. *Public Administration Review*, 1(4), 335-350.

Fitzgerald, L., Johnston, R., Brignall, T. J., Silvestro, R., & Voss, C. (1991). *Performance measurement in service businesses*. London: Chartered Institute of Management Accountants.

Fitzgerald, L., & Moon, P. (1996). *Performance measurement in service industries: making it work*, Chartered Institute of Management Accountants.

Folz, D. H. (2004). Service quality and benchmarking the performance of municipal services. *Public Administration Review*, 64(2), 209-220.

Forth, J., & McNabb, R. (2008). Workplace performance: A comparison of subjective and objective measures in the 2004 Workplace Employment Relations Survey. *Industrial Relations Journal*, 39(2), 104-123.

Fountain, J. J., & Roob, M. (1994). Service efforts and accomplishments measures: Development and use for government services. *Public Management*, 76(30), 6-12.

Franklin, A. (1999). Managing for results in Arizona: A fifth-year report card. *Public Productivity & Management Review*, 194-209.

Freel, M. S. (2003). Sectoral patterns of small firm innovation, networking and proximity. *Research Policy*, 32(5), 751-770.

Frederickson, H. G. (1980). *New public administration Tuscaloosa*, AL: University of Alabama Press.

Frederickson, H. G. (1990). Public administration and social equity. *Public Administration Review*, 50(2), 228-237.

Frederickson, H. G. (1996). Comparing the reinventing government movement with the New Public Administration. *Public Administration Review*, 56(3), 263-270.

Frederickson, H. G. (1997). *The spirit of public administration*: San Francisco, CA: Jossey-Bass.

Frederickson, H. G. (1999). Public ethics and the new managerialism. *Public Administration & Management*, 4(2), 299-324.

Frederickson, H. G. (2005). The state of social equity in American public administration. *National Civic Review*, 31–38.

Frederickson, H. G. (2010). *Social equity and public administration: Origins, developments, and applications*, Routledge.

Friedmann, L. (2002). The Microeconomics for Public Policy Analysis: Princeton NJ: Princeton University Press Weimer. In V. D. L., A. R. (2015) (Ed.), *Policy analysis: Concepts and practice* Routledge.

Friedrich, C. J. (1940). Public policy and the nature of administrative responsibility. In Carl Joachim Friedrich & Edward Sagendorph Mason (eds) (pp. 3-24). *Public Policy: A Yearbook of the Graduate School of Public Administration*: Cambridge, Harvard University Press.

Fry, B. R. (1989). *Mastering public administration,* Chatham, NJ: Chatham House.

Fryer, K., Antony, J., & Ogden, S. (2009). Performance management in the public sector. *International Journal of Public Sector Management*, 22(6), 478-498.

Furiady, O., & Kurnia, R. (2015). The effect of work experiences, competency, motivation, accountability and objectivity towards audit quality. *Procedia-Social and Behavioral Sciences*, 211, 328-335.

GAO. (1992). *Program performance measures: Federal agency collection and use of performance data*. (GAO/GGD-92-65). Washington, DC: U.S. Government Printing Office.

GAO. (2011). *Performance measurement and evaluation: Definitions and relationships*. (GAO 11-646SP). Washington, DC: U.S. Government Printing Office.

GAO. (2014). *Managing for results: Agencies' trends in the use of performance information to make decisions.* (GAO-14-747). Washington, DC: U.S. Government Printing Office

GAO. (2017). *Managing for results: Agencies' trends in the use of performance information to make decisions*. (GAO-17-775). Washington, DC: U.S. Government Printing Office

Ghotbabadi, A. R., Feiz, S., & Baharun, R. (2015). Service quality measurements: A review. *International Journal of Academic Research in Business and Social Sciences*, *5*(2), 267.

Giacchino, S., & Kakabadse, A. (2003). Successful policy implementation: the route to building self-confident government. *International Review of Administrative Sciences*, 69(2), 139-160.

Gibbs, R., Costa Lima, P., & Francozo, E. (2004). Metaphor is grounded in embodied experience. *Journal of Pragmatics,* 36(7), 1189-1210.

Gill, J., & Butler, R. J. (2003). Managing instability in cross-cultural alliances. *Long Range Planning*, 36(6), 543-563.

Gilly, J., & Torre, A. (2000). Proximity relations: Elements for an analytical framework. In M. R. B. e. Green M. B. (Ed.), *Industrial Networks and Proximity*: Ashgate Publishing, Aldershot.

Glick, W. H. (1985). Conceptualizing and measuring organizational and psychological climate: Pitfalls in multilevel research, *Academy of Management Review*, 10, 601-616.

Glick, W. H., & Roberts, K. H. (1984). Hypothesized interdependence, assumed independence. *Academy of Management Review*, 9(4), 722-735.

Goddard, R. D., Sweeland, S. R., & Hoy, W. K. (2000). Academic emphasis of urban elementary schools and student achievement in reading and mathematics: A multilevel analysis, *Educational Administration Quarterly*, 36(5), 683-702.

Gooden, S. (2014). *Race and Social Equity: A Nervous Area of Government.* Armonk, New York: M.E. Sharpe.

Government Performance and Results Act of 1993, Pub. L. 103-62, 107 Stat. In 103rd Congress, pp. 285-296. Retrieved from http://library.clerk.house.gov/reference-files/PPL_103_062_GovernmentPerformanceAndResultsAct_1993.pdf

Greenberg, J. (1990). Organizational justice: Yesterday, today, and tomorrow. *Journal of Management*, 16(2), 399-432.

Greenhaus, J. H., Parasuraman, S., & Mormley, W., M. (1990). Effects of race on organizational experiences, job performance evaluations, and career outcomes, *Academy of Management Journal*, 33(1), 64-86

Greiling, D. (2006). Performance measurement: a remedy for increasing the efficiency of public services? *International Journal of Productivity and Performance Management*, 55(6), 448-465.

Greunz, L. (2003). Geographically and technologically mediated knowledge spillovers between European regions. *Annals of Regional Science*, 37(4), 657–680.

Guy, M. E., & Newman, M. A. (2004). Women's Jobs, Men's Jobs: Sex Segregation and Emotional Labor. *Public Administration Review*, 289–298.

Hall, E. T. (1966). *The hidden dimension*. Anchor

Hamel, G., & Prahalad, C. K. (1989). Strategic intent. *Harvard Business Review*, 89, 63-76.

Hamilton, L. C. (2012). *Statistics with STATA: Version 12* (8th ed.) Cengage.

Hart, N., & Newcomer, K. (2018). *Presidential evidence initiatives: Lessons from the Bush and Obama administrations' efforts to improve government performance*. Biparitsan Policy Center

Harris, R. A., & Tichenor, D. J. (Eds.). (2009). *A History of the US Political System: Ideas, Interests, and Institutions [3 Volumes]: Ideas, Interests, and Institutions*. ABC-CLIO.

Harrison, X. A., Donaldson, L., Correa-Cano, M., Evans, J., Fisher, D. N., Goodwin, C. E. D., Robinson, B. S., Hodgson, D. J., & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology, *Peer J* 6: e4794; DOI 10.7717/peerj.4794

Hatry, H. P. (1997). We need a new concept of accountability, *The Public Manager: The New Bureaucrat*, 26(3), 37-39.

Hatry, H. P. (1999). Mini symposium on intergovernmental comparative performance data. *Public Administration Review*, 59(2), 101-102.

Hatry, H. P. (2002). Performance measurement: Fashions and fallacies. *Public Performance and Management Review*, 25(4), 352-358.

Hatry, H. P. (2006). *Performance measurement: Getting results*: The Urban Institute.

Hatry H. P., Winnie R. E., & Fisk D. M, (1973). *Practical program evaluation for state and local government officials*, The Urban Institute, Washington, DC.

Hedeker, D., & Gibbons, D. (2006*). Longitudinal data analysis*. Hoboken, NJ: John Wiley & Sons, Inc.

Heilizer, F. (1977). A review of theory and research on the assumptions of Miller's response competition (conflict) models: Response gradients. *The Journal of General Psychology*, 97(1), 17-71.

Hill, C. J., & Lynn Jr., L. E. (2009). *Public management: A three-dimensional approach*: Washington, DC: CQ Press.

Hess, D. B., & Almeida, T. M. (2007). Impact of proximity to light rail rapid transit on station-area property values in Buffalo, New York. *Urban Studies*, 44(5-6), 1041-1068.

Hofer, C. W. (1983). ROVA: A new measure for assessing organizational performance. *Advances in Strategic Management*, 2, 43-55

Hofer, C. W., & Schendel, D. (1978). *Strategy formulation: Analytical concepts*. St. Paul, MN: West.

Hofstede, G. (1980). Culture and organizations. *International Studies of Management & Organization*, 10(4), 15-41.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage publications

Hofstede, G., & Bond, M. H. (1988). The Confucius connection: From cultural roots to economic growth. *Organizational Dynamics*, 16(4), 5-21.

Hoggett, P. (2006). Conflict, ambivalence, and the contested purpose of public organizations. *Human Relations*, 59(2), 175-194.

Hollan, J., & Stornetta, S. (1992). Beyond being there. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 119-125). ACM.

Holmes, J. S., Amin Gutiérrez de Piñeres, S., & Douglas Kiel, L. (2006). Reforming government agencies internationally: is there a role for the balanced scorecard? *International Journal of Public Administration*, 29(12), 1125-1145.

Holzer, M., Charbonneau, E., & Kim, Y. (2009). Mapping the terrain of public service quality improvement: twenty-five years of trends and practices in the United States. *International Review of Administrative Sciences*, 75(3), 403-418.

Hood, C. (1991). A public management for all seasons? *Public Administration*, 69(1), 3-19.

Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Med*, 2000, 30, 1-15.

Horner, L., Lekhi, R., & Blaug, R. (2006). *Deliberative democracy and the role of public managers*: London: The Work Foundation.

Hox, J. (1998). Multilevel modeling: When and why. In I. Balderjahn, R. Mathar, & M. Schader (Eds.), *Classification, data analysis, and data highways* (pp. 147-154). Berlin, Germany: Springer.

Hsieh, J. Y. (2013). Impact of individual and organizational factors on job satisfaction: A comparison of multilevel models and multiple regression models using different data arrangements. *Journal of Management & Organization*, 19(1), 44-59.

Hu, H. H., Kandampully, J., & Juwaheer, T. D. (2009). Relationships and impacts of service quality, perceived value, customer satisfaction, and image: an empirical study. *The service Industries Journal*, 29(2), 111-125.

Huang, W. J., Chen, C. C., & Lin, Y. H. (2013). Cultural proximity and intention to visit: Destination image of Taiwan as perceived by Mainland Chinese visitors. *Journal of Destination Marketing & Management*, 2(3), 176-184.

Huber, G. P. & Power, D. J. (1985). Retrospective reports of strategic-level managers; Guidelines for increasing their accuracy, *Strategic Management Journal*, 6(2), 171-180

Hull, C. L. (1932). The goal-gradient hypothesis and maze learning. *Psychological Review*, 39(1), 25-43.

Hull, C. L. (1943). *Principles of behavior: An introduction to behavior theory*: Appleton-Century-Crofts, INC: New York.

Huta, V. (2014). When to use hierarchical linear modeling. *The Quantitative Methods for Psychology*, 10(1), 13-28.

IJzerman, H., & Semin, G. (2009). The thermometer of social relations: Mapping social proximity on temperature. *Psychological Science*, 20(10), 1214-1220.

Ingraham, P. W. (2005). Performance: Promises to keep and miles to go. *Public Administration Review*, 65(4), 390-395.

Jakobsen, M., & Jensen, R. (2015). Common method bias in public management studies. *International Public Management Journal*, 18(1), 3-30.

Jennings, E. T. (2005). Social equity and the Government Performance and Results Act. Paper presented at *the 8th Public Management Research Conference*, Los Angeles.

Jhang, J. H., & Lynch Jr, J. G. (2014). Pardon the interruption: Goal proximity, perceived spare time, and impatience. *Journal of Consumer Research*, 41(5), 1267-1283.

Johnson, A. (2001). Functions in innovation system approaches. Paper presented at the In *Nelson and Winter Conference*, Aalborg, Denmark.

Jones, T. M. (1991). Ethical decision making by individuals in organizations: An issue-contingent model. *The Academy of Management Review*, 16(2), 366-395.

Jones, L. R., & Kettl, D. (2003). Assessing public management reform in an international context. *International Public Management Review*, 4(1), 1-18.

Jørgensen, T. B., & Bozeman, B. (2007). Public values: An inventory. *Administration & Society*, 39(3), 354-381.

Julnes, P. D. L. (2006). Performance measurement: An effective tool for government accountability? The debate goes on, *Evaluation*, 12(2), 219-235.

Julnes, P. D. L. (2008). *Can Performance Measurement Support Program Performance Improvement and Accountability*? Thousand Oaks: Sage Publications, Inc.

Julnes, P. D. L., & Holzer, M. (2001). Promoting the utilization of performance measures in public organizations: An empirical study of factors affecting adoption and implementation, *Public Administration Review*, 61(6), 693-708

Jung, C. S. (2013). Organizational goal ambiguity and job satisfaction in the public sector. *Journal of Public Administration Research and Theory*, 24(4), 955-981.

Jung, C. S. & Lee, G. (2013). Goals, strategic planning, and performance in government agencies, *Public Management Review*, 15(6), 787-815.

Kamensky, J. M., & Abramson, M. A. (2002). Managing for results: Cutting-edge challenges facing government leaders in 2002. *Managing for results*, 3-18.

Kamensky, J. M., & Morales, A. (2005). *Managing for results*. Rowman & Littlefield Publishers, Incorporated.

Kaplan, R. S., & Norton, D. P. (1992). The Balanced Scorecard: Measures that Drive Performance. *Harvard Business Review*, 71-79.

Kaplan, R. S., & Norton, D. P. (2001). Transforming the balanced scorecard from performance measurement to strategic management: Part I. *Accounting Horizons*, 15(1), 87-104.

Kaplan, S. A., & Garrett, K. E. (2005). The use of logic models by community-based initiatives. *Evaluation and Program Planning*, 28(2), 167-172.

Kellough, J. E., & Naff, K. C. (2004). Responding to a wake-up call: An examination of federal agency diversity management programs. *Administration & Society*, 36(1), 62-90.

Kelly, G., Muers, S. & Mulgan, G., (2002) *Creating public value: An analytical framework for public service reform,* London, Cabinet Office

Kelly, J. M. (2005). The dilemma of the unsatisfied customer in a market model of public administration. *Public Administration Review*, 65(1), 76-84.

Kelly, J. M., & Swindell, D. (2002). Service quality variation across urban space: First steps toward a model of citizen satisfaction. *Journal of Urban Affairs*, 24(3), 271-288.

Kettl, D. (1994). *Reinventing Government? Appraising the National Performance Review*: Washington, DC: Brookings Institution.

Kettl, D. F. (2000). Public administration at the millennium: The state of the field. *Journal of Public Administration Research and Theory*, 10(1), 7-34.

Kiesler, S., & Sproull, L. (1982). Managerial response to changing environments: Perspectives on problem sensing from social cognition. *Administrative Science Quarterly*, 27, 548-570.

Kimberly, D. E., & Kramer, R. M. (2015). *Handbook of qualitative organizational research: Innovative pathways and methods*, Routledge.

Kimberly, J. R. (1980). Data aggregation in organization research: The temporal dimension, *Organization Studies*, 1(4), 367-377.

Kirat, T., & Lung, Y. (1999). Innovation and proximity – Territories as loci of collective learning processes. *European Urban and Regional Studies*, 6(1), 27-38.

Kirkman, B. L., Lowe, K. B., & Gibson, C. B. (2006). A quarter century of culture's consequences: A review of empirical research incorporating Hofstede's cultural values framework. *Journal of International Business Studies*, 37(3), 285-320.

Kirlin, J. J. (2001). Big Questions for a Significant Public Administration. *Public Administration Review*, 61(2), 140-143.

Kirschenbaum, D. S. (1985). Proximity and specificity of planning: A position paper. *Cognitive Therapy and Research*, 9(5), 489-506.

Kivetz, R., Urminsky, O., & Zheng, Y. (2006). The Goal-Gradient Hypothesis Resurrected: Purchase Acceleration, Illusionary Goal Progress, and Customer Retention. *Journal of Marketing Research*, 43(1), 39-58.

Klein, K. J., Conn, A. B., Smith, D. B., & Sorra, J. S. (2001). Is everyone in agreement? An exploration of within-group agreement in employee perceptions of the work environment. *Journal of Applied Psychology*, 86, 3-16

Klimoski, R., & Inks, L. (1990). Accountability forces in performance appraisal. *Organizational Behavior & Human Decision Processes*, 45(2), 194-208.

Knoben, J., & Oerlemans, L. A. G. (2006). Proximity and inter-organizational collaboration: A literature review. *International Journal of Management Reviews*, 8(2), 71-89.

Kobarg, S., Wollersheim, J. Welpe, I. M., & Spörrle, M. (2017). Individual ambidexterity and performance in the public sector: A multilevel analysis, *International Public Management Journal*, 20(2), 226-260

Koliba, C. J., Mills, R. M., & Zia, A. (2011). Accountability in governance networks: An assessment of public, private, and nonprofit emergency management practices following hurricane Katrina. *Public Administration Review*, 71(2), 210-220.

Koo, T, K. & Li, M. Y. (2016). A guideline of selecting and reporting Intraclass correlation coefficients for reliability research, *Journal of Chiropractic Medicine*, 15, 155-163

Koppell, J. (2008). Global governance organizations: Legitimacy and authority in conflict. *Journal of Public Administration Research and Theory*, 18(2), 177-203.

Kraut, R. E., Fussell, S. R., Brennan, S. E., & Siegel, J. (2002). Understanding effects of proximity on collaboration: Implications for technologies to support remote collaborative work. *Distributed Work*, 137-162.

Kreger, N. (2008). Proximal, distal, and the politics of causation: What's level got to do with it? *American Journal of Public Health*, 98(2), 221-230

Kull, D. (1978). Productivity Programs in Federal Government. *Public Administration Review*, 38(1), 5-9.

Lackey, K. J., & Kaczynski, A. T. (2009). Correspondence of perceived vs. objective proximity to parks and their relationship to park-based physical activity. *International Journal of Behavioral Nutrition and Physical Activity*, 6(1), 53.

Lam, A. (1997). Embedded firms, embedded knowledge: Problems of collaboration and knowledge transfer in global cooperative ventures. *Organization Studies*, 18(6), 973-996.

Lan, Z. & Rainey, H. G. (1992). Goals, rules, and effectiveness in public private, and hybrid organizations: More evidence on frequent assertions about differences, *Journal of Public Administration Research and Theory*, 2(1), 5-28.

Lane, P. J., & Lubatkin, M. (1998). Relative absorptive capacity and interorganizational learning. *Strategic Management Journal*, 19(5), 461-477.

Langbein, L. I. (1980). Goals and criteria for evaluating public programs. In *Discovering whether programs work: A guide to statistical methods for program evaluation*: Santa Monica, CA: Goodyear Publishing Co.

Langley, A. (1999). Strategies for theorizing from process data. *Academy of Management Review*, 24(4), 691-710

Lasswell, H. D. (2018). *Politics: Who gets what, when, how*. Pickle Partners Publishing.

Lawler, E. (2003). Reward practices and performance management. *Organizational Dynamics*, 32(4), 396-404.

Leary, T., Ridley, S., Burchett, K., Kong, A., Crispin, P., & Wright, M. (2002). Assessing critical care unit performance: A global measure using graphical analysis. *Anaesthesia*, 57, 751–755

Lee, K. M. (2004). Presence, explicated. *Communication Theory*, 14(1), 27-50.

Lee, R., & Wilbur, E., R. (1985). Age, education, job tenure, salary, job characteristics, and job satisfaction: A multivariate analysis, *Human Relations*, 38(8), 781-791.

Lent, L. V., Sungur, H., Kunneman, F. A., & Das, E. (2017). Too far to care? The role of psychological proximity in public attention and fear for Ebola. *J Med Internet Res*, 19(6), e193 doi:10.2196/jmir.7219

Lewis, F. L. (1991). *Introduction to Total Quality Management in the federal government*, Washington, DC: Federal Quality Institute.

Lin, Y. T., Chang, T. C., & Chen, K. S. (2018). A novel approach to evaluating the performance of physical fitness by combining statistical inference with the Radar Chart. *Journal of Testing and Evaluation*, 46(4), 1498-1507.

Linna, A., Elovainio, M., Van den Bos, K., Kivimäki, M., Pentti, J., & Vahtera, J. (2012). Can usefulness of performance appraisal interviews change organizational justice perceptions? A

4-year longitudinal study among public sector employees. *The International Journal of Human Resource Management,* 23(7), 1360-1375.

Lipsky, M. (1980). *Street-level bureaucracy: Dilemmas of the individual in public services*: Russell Sage Foundation.

Liu, B., Yang, K., & Yu, W. (2015). Work-related stressors and health-related outcomes in public service: Examining the role of public service motivation, *The American Review of Public Administration*, 45(6), 653-673.

Ljungholm, D. P. (2015). The impact of transparency in enhancing public sector performance. *Contemporary Readings in Law and Social Justice,* 7(1), 172-178.

Lombard, M., Reich, R., Grabe, M., Bracken, C., & Ditton, T. (2000). Presence and television: The role of screen size. *Human Communication Research*, 26, 75–98.

Lundvall, B. A. (1993). Explaining interfirm cooperation and innovation. Limits of the transaction-cost approach. In (pp. 52-64). Grabher, G. (Ed) *The embedded firm: On the socioeconomics of industrial networks*: Routledge, London.

Lynn Jr., L. E. (2006). *Public management: Old and new*: Routledge.

Lynn Jr, L. E., Heinrich, C. J., & Hill, C. J. (2000). Studying governance and public management: Challenges and prospects. *Journal of Public Administration Research and Theory*, 10(2), 233-262.

Lynn Jr., L. E., Heinrich, C. J., & Hill, C. J. (2001). *Improving governance: A new logic for research*. Washington, DC: Georgetown University Press.

MacKinnon, D., Cumbers, A., & Chapman, K. (2002). Learning, innovation, and regional development: a critical appraisal of recent debates. *Progress in Human Geography,* 26(3), 293-311.

Malecki, E. J., & Oinas, P. (1999). *Making connections. Technological learning and regional economic change*: Ashgate, Aldershot.

Martin, L. L., & Tesser, A. (1996). *Striving and feeling: Interactions among goals, affect, and self-regulation*: Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Martins, L. L., Gilson, L. L., & Maynard, M. T. (2004). Virtual teams: What do we know and where do we go from here? *Journal of Management*, 30(6), 805-835.

Maskell, P. (2001). Towards a knowledge-based theory of the geographical cluster. *Industrial and Corporate Change*, 10(4), 921-943.

Maas, C. J. M, & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling, *Methodology*, 1(3), 59-92

Matthews, J. L., & Matlock, T. (2011). Understanding the link between spatial distance and social distance. *Social Psychology*, 42(3), 1-24

McDougall, G. H., & Levesque, T. (2000). Customer satisfaction with services: putting perceived value into the equation. *Journal of Services Marketing*, 14(5), 392-410.

McGraw, K., & Wong, S. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1(1), 30–46.

McKay, P. F., Avery, D. R., & Morris, M. A. (2009). A tale of two climates: diversity climate from subordinates' and managers' perspectives and their role in store unit sales performance. *Personnel Psychology, 62*(4), 767-791.

Meier, K. J., & O'Toole Jr, L. J. (2006). *Bureaucracy in a democratic state: A governance perspective*: Johns Hopkins University Press.

Meier, K. J., & O'Toole, L. J. (2012). Subjective organizational performance and measurement error: Common source bias and spurious relationships. *Journal of Public Administration Research and Theory*, 23(2), 429-456.

Meier, K.J. & L.J. O'Toole Jr. (2013a). I think (I Am Doing Well), therefore I am: Assessing the validity of administrators' self-assessments of performance, *International Public Management Journal*, 16(1), 1–27.

Meier, K. J. & O'Toole, L. J. (2013b). Subjective organizational performance and measurement error: Common source bias and spurious relationships, *Journal of Public Administration Research and Theory*, 23, 429-456

Meisters, C., & Werker, C. (2004). Physical and organizational proximity in territorial innovation systems: Introduction to the special issue. *Journal of Economic Geography*, 4(1), 1-2.

Melkers, J., & Willoughby, K. (2005). Models of performance-measurement use in local governments: Understanding budgeting, communication, and lasting effects. *Public Administration Review,* 65(2), 180-190.

Mencl, J., & May, D. R. (2009). The effects of proximity and empathy on ethical decision-making: An exploratory investigation. *Journal of Business Ethics*, 85(2), 201-226.

Merget, A. E., & Renee, A. B. (1982). Equity as a decision rule in local services. In R. C. Rich (Ed.), *Analyzing urban service distributions* (pp. 21-44): Lexington, MA: D.C. Heath.

Mero, N. P., Guidice, R. M., & Anna, A. L. (2006). The interacting effect of accountability and individual differences on rater response to a performance-rating task. *Journal of Applied Social Psychology,* 36(4), 795-819.

Meyer, J. P., & Allen, N. J. (1997). *Commitment in the workplace: Theory, research, and application*, SAGE Publications

Meyer, M. M. 2002. *Rethinking performance measurement: Beyond the balanced scorecard*. Cambridge, UK: Cambridge University Press.

Meyners, J., Barrot, C., Becker, J. U., & Goldenberg, J. (2017). The role of mere closeness: How geographic proximity affects social influence, *Journal of Marketing*, 81(5), 49-66.

Milakovich, M. (1995). *Improving service quality: Achieving high performance in the public and private sectors*. Delray Beach, FL: St. Lucie Press.

Miller, G. J., & Moe, T. M. (1983). Bureaucrats, legislators, and the size of government. *American Political Science Review*, 77(2), 297-322.

Mitchell, F., Nielsen, L. B., Nørreklit, H., & Nørreklit, L. (2013). Scoring strategic performance: a pragmatic constructivist approach to strategic performance measurement, *Journal of Management & Governance*, 17(1), 5-34.

Moe, R. C., & Gilmour, R. S. (1995). Rediscovering principles of public administration: The neglected foundation of public law. *Public Administration Review*, 55(2), 135-146.

Mokhlis, S. A. (2011**)** Municipal Service Qualityand Citizen Satisfaction in Southern Thailand, *Journal of Public Administration and Governance*, 1, 122-137.

Monge, P., & Kirste, K. (1980). Measuring proximity in human organization. *Social Psychology Quarterly*, 43(1), 110-115.

Moore, M. H. (1995). *Creating public value: Strategic management in government*: Harvard university press.

Moore, M. H. (2012). *Recognizing public value: developing a public value account and a public value scorecard*. John F. Kennedy School of Government, Harvard University. Cambridge, MA. USA.

Morgan, D. L. (1993). Qualitative content analysis: A guide to paths not taken, *Qualitative Health Research*, 3(1), 112-121.

Morgan, J. (1991). From artifacts to archives: Digging into a community's past. *Journal of Geography*, 90 (4): 179–181.

Mosher, F. C. (1981). *Basic literature of American public administration*, 1787-1950: Holmes & Meier Pub.

Mosher, F. C., & Stillman, R. J. (1982). *Professions in government*. Transaction Pub.

Mosley, H. and Mayer, A. (1999). Benchmarking national labour market performance: A radar chart approach, discussion paper, *Social science research centre*, Berlin, ISSN Nr. Pp. 1011-9523. FS I 99-202

Mowen, A., Orsega-Smith, E., Payne, L., Ainsworth, B., & Godbey, G. (2007). The role of park proximity and social support in shaping park visitation, physical activity, and perceived health among older adults. *Journal of Physical Activity & Health*, 4(2), 167-179.

Moynihan, D. P. (2008a). *The dynamics of performance management: Constructing information and reform*: Washington, DC: Georgetown University Press.

Moynihan, D. (2008b). Advocacy and learning: An interactive-dialogue approach to performance information use. In *Performance information in the Public Sector* (pp. 24-41). Palgrave Macmillan, London.

Moynihan, D. P., & Lavertu, S. (2012). Does involvement in performance management routines encourage performance information use? Evaluating GPRA and PART. *Public Administration Review*, 72(4), 592-602.

Moynihan, D. P., & Pandey, S. K. (2004). Testing how management matters in an era of government by performance management. *Journal of Public Administration Research and Theory*, 15(3), 421-439.

Mulvaney, M. (2018) Memorandum for heads of executive departments and agencies: Improving the management of federal programs and projects through implementing the Program Management Improvement Accountability Act (PMIAA), *Office of Management and Budget*, June, 25

Murphy, K. R., & Cleveland, J. (1995). *Understanding performance appraisal: Social, organizational, and goal-based perspectives*: Sage.

Nabatchi, T. (2012). Putting the "public" back in public values research: Designing participation to identify and respond to values. *Public Administration Review*, 72(5), 699-708.

Neugerg, S. L. (1989). The goal of forming accurate impressions during social interactions: Attenuating the impact of negative expectancies, *Journal of Personality and Social Psychology*, 56(3), 374-386.

Newcomb, T. M. (1956). The prediction of interpersonal attraction. *American Psychologist*, 11(11), 575.

Newcomer, K. (2007). How does Program Performance Assessment Affect Program Management in the Federal Government? *Public Performance & Management Review*, 30(3), 332-350.

Niskanen, W. A. (1971). *Bureaucracy and Representative Government*: Transaction Publishers.

Noordegraaf, M. (2007). From “pure” to “hybrid” professionalism: Present-day professionalism in ambiguous public domains. *Administration & Society*, 39(6), 761-785.

Nooteboom, B. (2000). Learning by interaction: absorptive capacity, cognitive distance and governance. *Journal of Management and Governance*, 4(1-2), 69-92.

Norman-Major, K. (2011). Balancing the four Es; or Can we achieve equity for social equity in public administration? *Journal of Public Affairs Education*, 17(2), 233-252.

Nunnally, J. C. (1978). *Psychometric Theory*. McGraw-Hill Book Company, pp. 86-113, 190-255.

Nyhan, R. C., & Martin, L. L. (1999). Comparative performance measurement: A primer on data envelopment analysis. *Public Productivity & Management Review*, 348-364.

OMB (2017) Section 200—Overview of the federal performance framework, OMB Circular No. A-11

Oerlemans, L., & Meeus, M. (2005). Do organizational and spatial proximity impact on firm performance? *Regional Studies*, 39(1), 89-104.

Oerlemans, L., Meeus, M., & Boekema, F. (2001). On the spatial embeddedness of innovation networks: an exploration of the proximity effect. *Tijdschrift voor economische en sociale geografie*, 92(1), 60-75.

O'Flynn, J. (2007). From new public management to public value: Paradigmatic change and managerial implications. *Australian Journal of Public Administration*, 66(3), 353-366.

O'Leary, M. B., Wilson, J. M., & Metiu, A. (2014). Beyond being there: The symbolic role of communication and identification in perceptions of proximity to geographically dispersed colleagues. *MIS quarterly*, 38(4), 1219-1243.

Olson, J. S., Teasley, S., Covi, L., & Olson, G. (2002). The (currently) unique advantages of collocated work. *Distributed Work,* 113-135.

Ömürgönülşen, U. (2002). Performance Measurement in the Public Sector: Rising Concern, Problems in Practice and Prospects. *Hacettepe Üniversitesi İİBF Dergisi*, 20(1), 99-134.

Onwuegbuzie, A. J., Johnson, R. B., & Collins, K. M. (2009). Call for mixed analysis: A philosophical framework for combining qualitative and quantitative approaches. International *Journal of Multiple Research Approaches*, 3(2), 114-139.

Osborne, D., & Gaebler, T. (1992). *Reinventing government*, New York, NY: Penguin Press.

Ossege, C. (2012). Accountability–are We Better off Without It? An empirical study on the effects of accountability on public managers' work behaviour. *Public Management Review*, *14*(5), 585-607.

Ostroff, C., & Schmitt, N. (1993). Configurations of organizational effectiveness and efficiency. *Academy of Management Journal,* 36(6), 1345-1361.

Otley, D. (1999). Performance management: a framework for management control systems research. *Management Accounting Research*, 10(4), 363-382.

O'Toole Jr, L. J., & Meier, K. J. (2015). Public management, context, and performance: In quest of a more general theory. *Journal of Public Administration Research and Theory*, 25(1), 237-256.

Otrusinova, M. & Pastuszkova, E. (2012). Concept of 3E's and public administration performance, *International Journal of Systems Applications, Engineering & Development*, 6(2), 171-178.

Pandey, S. K., & Rainey, H. G. (2006). Public managers' perceptions of organizational goal ambiguity: Analyzing alternative models. *International Public Management Journal*, 9(2), 85-112.

Patil, S.V., Vieider, F. & Tetlock, P.E. (2014). Process versus outcome accountability. *In M. Bovens, R.E. Goodin & T. Schillemans (Eds.), The Oxford Handbook of Public Accountability. New York: Oxford University Press*

Pettigrew, A. M. (1979). On Studying Organizational Cultures Andrew M. Pettigrew. *Administrative Science Quarterly*, 24(4), 570-581.

Piotrowski, S. J., & Rosenbloom, D. H. (2002). Nonmission-based values in results-oriented public management: The case of freedom of information. *Public Administration Review*, 62(6), 643-657.

Pitts, D. W. (2005). Diversity, representation, and performance: Evidence about race and ethnicity in public organizations. *Journal of Public Administration Research and Theory*, 15(4), 615-631.

Pitts, D. W., Hicklin, A. K., Hawes, D. P., & Melton, E. (2010). What drives the implementation of diversity management programs? Evidence from public organizations. *Journal of Public Administration Research and Theory*, 20(4), 867-886.

Poister, T. H. (2008). *Measuring performance in public and nonprofit organizations*: Jossey-Bass.

Poister, T. H., & Streib, G. (1999). Performance Measurement in Municipal Government: Assessing the State of the Practice. *Public Administration Review*, 59(4), 325-335.

Pollanen, R. M. (2014). Annual performance reporting as accountability mechanism in local government, *International Journal of Business and Public Administration*, 11(2)

Pollitt, C. (2002). Clarifying convergence: Striking similarities and durable differences in public management reform. *Public Management Review*, 4(1), 471-492.

Pollitt, C., & Bouckaert, G. (2000). *Public management reform: An international comparison*. Oxford, UK: Oxford University Press.

Ponds, R., Van Oort, F., & Frenken, K. (2007). The geographical and institutional proximity of research collaboration. *Papers in Regional Science*, *86*(3), 423-443.

Porter, T. (1995). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton, NJ: Princeton University Press.

Quinn, K. M. (2004). Bayesian factor analysis for mixed ordinal and continuous responses. *Political Analysis*, 12(4), 338-353.

Quinn, R. E., & Rohrbaugh, J. (1983). A spatial model of effectiveness criteria: Towards a competing values approach to organizational analysis. *Management Science*, 29(3), 363–377.

Radin, B. (2006). C*hallenging the performance movement: Accountability, complexity, and democratic values*: Georgetown University Press.

Rainey, H. G. (1989). Public management: Recent research on the political context and managerial roles, structures and behaviors. *Journal of Management*, 15(2), 229–250.

Rainey, H. G. (2014). *Understanding and managing public organizations* (5th ed.): Jossey-Bass

Rainey, H. G., & Steinbauer, P. (1999). Galloping Elephants: Developing Elements of a Theory of Effective Government Organizations. *Journal of Public Administration Research and Theory*, 9(1), 1-32.

Rajesh, R., Pugazhendhi, S., Ganesh, K. and Muralidharan, C. (2010). Service dimension based descriptive study of third-party logistics (3PL) service provider and 3PL users, *International Journal of Decision Making in Supply Chain and Logistics*, 1(1), 1-8.

Raudenbush, S. W., & Bryk, A. S. (2001). *Hierarchical models: Applications and data analysis methods*. (2nd ed.). SAGE Publications, Inc.

Ravindran, A. R. (2007). *Operations research and management science handbook*: CRC Press

Reed, J. A., & Phillips, D. A. (2005). Relationships between physical activity and the proximity of exercise facilities and home exercise equipment used by undergraduate university students. *Journal of American College Health*, 53(6), 285-290.

Rendon, R. G. (2008). Procurement process maturity: Key to performance measurement. *Journal of Public Procurement*. 8(2), 200-214.

Rice, R. E., & Aydin, C. (1991). Attitudes toward new organizational technology: Network proximity as a mechanism for social information processing. *Administrative Science Quarterly*, 219-244.

Rich, R. C. (1979). Neglected issues in the study of urban service distributions: A research agenda. *Urban Studies,* 16(2), 143-156.

Ridley, C. E., & Simon, H. A. (1943). *Measuring municipal activities* (2nd ed.): Chicago: International City Managers' Association.

Riege, A., & Lindsay, N. (2006). Knowledge management in the public sector: stakeholder partnerships in the public policy development, *Journal of Knowledge Management*, 10(3), 24-39.

Rivera, M., Soderstrom, S., & Uzzi, B. (2010). Dynamics of dyads in social networks: Assortative, relational, and proximity mechanisms. *Annual Review of Sociology*, 36, 91-115.

Rolle, A., & Liu, K. (2007). An empirical analysis of horizontal and vertical equity in the public schools of Tennessee, 1994-2003. *Journal of Education Finance*, 328-351.

Rosenbaum, D., Schuck, A., Graziano, L., & Stephens, C. (2007). Measuring police and community performance using web-based surveys: *Findings from the Chicago internet project: Center for Research in Law and Justice*.

Rosenbloom, D. H. (1983). Public administrative theory and the separation of powers. *Public Administration Review*, 43(3), 219-227.

Rourke, F. E. (1992). Responsiveness and neutral competence in American bureaucracy. *Public Administration Review*, 539-546.

Ryan, R. M., & Deci, E. L. (2006). Self-regulation and the problem of human autonomy: Does psychology need choice, self-determination, and will? *Journal of Personality, 74*(6), 1557-1586.

Sabharwal, M. (2014). Is diversity management sufficient? Organizational inclusion to further performance. *Public Personnel Management*, 43(2), 197-217.

Salazar, J. M., & Marin, G. (1977). National stereotypes as a function of conflict and territorial proximity: A test of the mirror image hypothesis. *The Journal of Social Psychology*, 101(1), 13-19.

Saltzstein, Grace H. 1992. Bureaucratic responsiveness: Conceptual issues and current research. *Journal of Public Administration Research and Theory* 2(1): 63-88.

Sanderson, I. (2001). Performance management, evaluation and learning in 'modern' local government. *Public Administration*, 79(2), 297-313.

Sandfort, J. R. (2000). Moving beyond discretion and outcomes: Examining public management from the front lines of the welfare system. *Journal of Public Administration Research and Theory*, 10(4), 729-756.

Sanger, M. B. (2008). From measurement to management: Breaking through the barriers to state and local performance. *Public Administration Review*, 68, S70-S85.

Schamp, E. W., Rentmeister, B., & Lo, V. (2004). Dimensions of proximity in knowledge-based networks: the cases of investment banking and automobile design. *European Planning Studies*, 12(5), 607-624.

Schermerhorn, J. R., Osborn, R., & Hunt, J. G. (2007). *Organizational behavior*: New York: Wiley.

Schick, A. (2001). Getting performance measures to measure up. In *Quicker, better, cheaper? Managing performance in American government*, 39-60.

Schneider, B. (1987). The people make the place. *Personnel Psychology*, *40*(3), 437-453.

Scott, C. (1999). Communication technology and group communication. In *Handbook of group communication theory and research*. L.R. Frey, D. Gouran and M. S. Poole (eds), 432- 474, Thousand Oaks, CA: Sage.

Scutaru, C., Quarcoo, D., Takemura, M., Welte, T., Fischer, T. C., & Groneberg-Kloft, B. (2010). Density-equalizing mapping and scientometric benchmarking in Industrial Health. *Industrial health*, 48(2), 197-203.

Seiler, M. J., Seiler, Vi.L., Traub, S., & Harrison, D. M. (2008). Familiarity bias and the status quo alternative, *Journal of Housing Research*, 17(2), 139-154.

Selden, S. C., Brewer, G. A., & Brudney, J. L. (1999). Reconciling competing values in public administration: Understanding the administrative role concept. *Administration & Society*, 31(2), 171-204.

Sezhian, M. V., Muralidharan, C., Nambirajan, T., & Deshmukh, S. G. (2011). Developing a performance importance matrix for a public sector bus transport company: A case study, *Theoretical and Empirical Researches in Urban Management*, 6(3), 5-14.

Shepperd, J. A., Ouellette, J. A., & Fernandez, J. K. (1996). Abandoning unrealistic optimism: Performance estimates and the temporal proximity of self-relevant feedback. *Journal of Personality and Social Psychology*, 70(4), 844.

Shieh, G. (2016). Choosing the best index for the average score intraclass correlation coefficient, *Behavior Research Meth*ods, 48(3), 994-1003.

Shinkle, G. A. (2012). Organizational aspirations, reference points, and goals: Building on the past and aiming for the future, *Journal of Management*, 38(1), 415-455

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. John Wiley & Sons.

Simon, H. A. (1946). The proverbs of administration. *Classics of Organization Theory*, 3, 101-113.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63(2), 129.

Smith-Lovin, L. (2003). Self, identity, and interaction in an ecology of identities. In *Advances in identity theory and research* (pp. 167-178). Springer, Boston, MA.

Smith, G. E., & Huntsman, C. A. (1997). Reframing the metaphor of the citizen-government relationship: A value-centered perspective. *Public Administration Review*, 57(4), 309-318.

Smith, P. (1996). *Measuring outcomes in the public sector*: Taylor & Francis, London.

Sofyani, H., Riyadh, H. A., & Fahlevi, H. (2020). Improving service quality, accountability and transparency of local government: The intervening role of information technology governance. *Cogent Business & Management*, 7(1).

Sommer, R. (1962). The distance for comfortable conversation: A further study. *Sociometry*, 25(1), 111-116.

Sorber, B. (1993). Performance measurement in the central government departments of the Netherlands. *Public Productivity & Management Review*, 17(1), 59-68.

Sowa, J. E., & Selden, S. C. (2003). Administrative discretion and active representation: An expansion of the theory of representative bureaucracy. *Public Administration Review*, 63(6), 700-710.

Starbuck, W. H., & Mezias, J. M. (1996). Opening Pandora's box: Studying the accuracy of managers' perceptions, *Journal of Organizational Behavior*, 17(2), 99-117.

Starling, G. (2010). *Managing the public sector*: Cengage Learning.

Stern, C., Cole, S., Gollwitzer, P., Oettingen, G., & Balcetis, E. (2013). Effects of implementation intentions on anxiety, perceived proximity, and motor performance. *Personality and Social Psychology Bulletin*, 39(5), 623-635.

Sternberg, R. (1999). Innovative linkages and proximity: Empirical results from recent surveys of small and medium-sized firms in German regions. *Regional Studies*, 33(6), 529-540.

Stoker, G. (2006). Public value management: A new narrative for networked governance? *Public Administration*, 36(1), 41-57.

Stone, D. C. (2002). *Policy paradox: The art of political decision making*, revised edition: London and New York, NY: WW Norton and Company.

Svara, J. H., & Brunet, J. R. (2004). Filling in the skeletal pillar: Addressing social equity in introductory courses in public administration. *Journal of Public Affairs Education*, 10(2), 99-109.

Svara, J. H., & Brunet, J. R. (2005). Social equity is a pillar of Public Administration. *Journal of Public Affairs Education*, 253-258.

Swiss, J. E. (1992). Adapting Total Quality Management (TQM) to government. *Public Administration Review*, 52(4): 426-438.

Talbot, C. (1999). Public performance-towards a new model? *Public Policy and Administration*, 14(3), 15-34.

Talbot, C. (2010). *Theories of performance: organizational and service improvement in the public domain:* Oxford University Press Inc., New York.

Taylor, F. W. (1911). *The principles of scientific management*: Harper & Brothers: New York and London.

Tenenbaum, G., Pinchas, S., Elbaz, G., Bar-Eli, M., & Weinberg, R. (1991). Effect of goal proximity and goal specificity on muscular endurance performance: A replication and extension. *Journal of Sport and Exercise Psychology*, 13(2), 174-187.

The Performance Institute. (2018). Five things you should know about Trump's 'President's Management Agenda. Retrieved from https://www.performanceinstitute.org/2017/03/16/5-things-know-trumps-new-presidents- management-agenda/

Tigue, P., & Strachota, D. (1994). *The use of performance measures in city and county budgets*. Chicago: Government Finance Officer Association.

Torre, A., & Rallet, A. (2005). Proximity and localization. *Regional Studies*, 39(1), 47- 59.

Torre, A., & Zuindeau, B. (2009). Proximity economics and environment: Assessment and prospects. *Journal of Environmental Planning and Management*, 52(1), 1-24.

Tremblay, D. G., Klein, J. L., Fontan, J. M., & Rousseau, S. (2003). Territorial proximity and innovation: a survey of the Montreal region. *Revue d'Economie Regionale et Urbaine*, 5, 835-852.

Tresch, R. W. (2008). *Public sector economics*: Palgrave Macmillan.

Trodden, S. A. (1996). The objectives and performance of performance auditing: The perspective of a United States, in *Performance auditing and the modernization of government*, Organization for Economics

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, 117(2), 440-463.

Tse, D. K., & Wilton, P. C. (1988). Models of consumer satisfaction formation: An extension, *Journal of Marketing Research*, 25(2), 204-212.

Tushman, M. L., & Anderson, P. (1986). Technological discontinuities and organizational environments. *Administrative Science Quarterly*, 31(3), 439-465.

U.S. Senate. Committee on Governmental Affairs. 1993. Report on the Government Performance and Results Act of 1993. http://www.whitehouse.gov/omb/mgmt-gpra/gprptm

Urry, J. (2002). Mobility and proximity. *Sociology*, 36(2), 255-274.

Uzzi, B. (1997). Social structure and competition in interfirm networks: the paradox of embeddedness. *Administrative Science Quarterly* 42(1), 35-67.

Van de Walle, S., & Van Dooren, W. (2010). How is information used to improve performance in the public sector: Exploring the dynamics of performance information? In: *Walshe, K.*,

Harvey, G. & Jas, P. (eds). Connecting knowledge and performance in public services: From knowing to doing. Cambridge: Cambridge University Press, pp. 33-54.

Van Dooren, W., Bouckaert, G., & Halligan, J. (2010). *Performance management in the public sector*. Routledge.

Van Dooren, W., & Van de Walle, S. (2008). *Performance information in the public sector: How it is used*, Palgrave Macmillan

Van Mierlo, H., Vermunt, J. K., & Rutte, C. G. (2009). Composing group-level constructs from individual-level survey data. *Organizational Research Methods*, 12, 368-392.

Vermeeren, B., Kuipers, B., & Steijin, B. (2011). Two faces of the satisfaction mirror: A study of work environment, job satisfaction, and customer satisfaction in Dutch municipalities, *Review of Public Personnel Administration*, 31(2), 171-189.

Wagenheim, G. D. & Reurink, J. H. (1991). *Public Administration Review*, 51(3), 263-270.

Waldo, D. (1948). *The administrative state*. The Ronald Press Company: New York.

Waldo, D. (1968a). Public administration. *Journal of Politics*, *30*(2), 443-479.

Walker, R. M., Boyne, G. A., & Brewer, G. A. (2010). *Public management and performance: Research directions*. Cambridge University Press.

Walker, R. M., & Brewer, G. A. (2008). An organizational echelon analysis of the determinants of red tape in public organizations. *Public Administration Review*, 68(6), (November/December): 1112-1127.

Walker, R. M., Brewer, G. A., Boyne, G. A., & Avellaneda, C. N. (2011). Market orientation and public service performance: New public management gone mad? *Public Administration Review*, 71(5), 707-717.

Wall, T., Michie, J., Patterson, M., Wood S., Sheehan, M., Clegg, C. W. & West, M (2004), On the validity of subjective measures of company performance, *Personnel Psychology*, 57(1), 95-118.

Wang, Q. (2016). Remembering the self in cultural contexts: A cultural dynamic theory of autobiographical memory. *Memory Studies*, 9(3), 295-304.

Wang, X. (2002). Assessing administrative accountability: Results from a national survey. *The American Review of Public Administration*, 32(3), 350-370.

Wang, X. & Giannakis, A. G. (1999). Public officials' attitudes toward subjective performance measures, *Public Productivity & Management Review*, 22(4), 537-553.

Weimer, D., & Vining, A. (2015). *Policy Analysis: Concepts and practice*: Routledge.

Weiner, J. (2007). *Measurement: reliability and validity measures*. Baltimore: John Hopkins Bloomberg, School of Public Health.

Weiss, J. A., & Piderit, S. K. (1999). The value of mission statements in public agencies. *Journal of Public Administration Research and Theory,* 9(2), 193-224.

Wellen, J., & Neale, M. (2006). Deviance, self-typicality, and group cohesion: The corrosive effects of the bad apples on the barrel. *Small Group Research*, 37(2), 165-186.

Whitford, A. B., Lee, S.-Y., Yun, T., & Jung, C. S. (2010). Collaborative behavior and the performance of government agencies. *International Public Management Journal*, 13(4), 321-349.

WhiteHouse.gov. (2017). Section 230-Agency strategic planning.

Wholey, J. S. (1999). Performance-based management: Responding to the challenges. *Public Productivity & Management Review*, 22(3), 288-307.

Wholey, J. S., & Hatry, H. P. (1992). The case for performance monitoring. *Public Administration Review*, 52(6), 604-610.

Wholey, J. S., Hatry, H. P., & Newcomer, K. E. (2010). *Handbook of practical program evaluation*, Jossey-Bass.

Wilhelm, M., & Ritz, B. (2003). Residential proximity to traffic and adverse birth outcomes in Los Angeles County, California 1994-1996. *Environmental Health Perspectives*, 111(2), 207-216.

Wilkof, M. V., Brown, D. W., & Selsky, J. W. (1995). When the stories are different: The influence of corporate culture mismatches on interorganizational relations. *The Journal of Applied Behavioral Science*, 31, 373-388.

Williams, D. W. (2002). Before performance measurement. *Administrative Theory and Praxis*, 24(6), 457-486.

Williams, D. W. (2003). Measuring government in the early twentieth century. *Public Administration Review,* 63(6), 643-659.

Williams, L. E., Stein, R., & Galguera, L. (2014). The distinct affective consequences of psychological distance and construal level. *Journal of Consumer Research*, 40(6), 1123-1138.

Wilson, J. M., O'Leary, M. B., Metiu, A., & Jett, Q. R. (2008). Perceived proximity in virtual work: Explaining the paradox of far-but-close. *Organization Studies* 29(7), 979-1002.

Winiewski, M. and Donelly, M. (1996) Measurin Service Quality in The Public Sector: The Potential for SERVQUAL, *Total Quality Management Journal*, 7, 357-365.

Witt, J. K. (2011). Action's effect on perception. *Current Directions in Psychological Science*, 20, 201-206.

Witt, J. K., & Dorsch, T. E. (2009). Kicking to bigger uprights: Field goal kicking performance influences perceived size. *Perception*, 38(9), 1328-1340.

Witt, J. K., Linkenauger, S. A., Bakdash, J. Z., & Proffitt, D. R. (2008). Putting to a bigger hole: Golf performance relates to perceived size. *Psychonomic Bulletin & Review*, 15(3), 581-585.

Witt, J. K., Linkenauger, S. A., & Proffitt, D. R. (2012). Get me out of this slump! Visual illusions improve sports performance. *Psychological Science*, *23*(4), 397-399.

Wooldridge, B., & Gooden, S. (2009). The epic of social equity: Evolution, essence and emergence. *Administrative Theory and Praxis*, 31(2), 222-234.

Wray, L., & Hauer, J. (1997). Performance measurement to achieve quality of life: Adding value through citizens. *Public Management*, 79.

Wuyts, S., Colombo, M. G., Dutta, S., & Nooteboom, B. (2005). Empirical tests of optimal cognitive distance. *Journal of Economic Behaviour & Organization* 58(2), 277-230.

Wynder, M. (2014). The effect of physical proximity on investment in corporate social responsibility initiatives: A exploratory analysis. *e-Journal of Social & Behavioural Research in Business*, 5(1), 13-24.

Yang, K. (2009). Examining perceived honest performance reporting by public organizations: Bureaucratic politics and organizational practice, *Journal of Public Administration Research and Theory*, 19 (1), 81–105.

Yang, K., & Callahan, K. (2007). Citizen involvement efforts and bureaucratic responsiveness: Participatory values, stakeholder pressures, and administrative practicality. *Public Administration Review*, 67(2), 249-264.

Yang, K., & Holzer, M. (2006). The performance-trust link: Implications for performance measurement. *Public Administration Review*, 66(1), 114-126.

Yang, K., & Hsieh, J. Y. (2007). Managerial effectiveness of government performance measurement: Testing a middle-range model, *Public Administration Review*, 67(5), 861-879

Yang, K., & Rho, S. Y. (2007). E-government for better performance: Promises, realities, and challenges. *International Journal of Public Administration*, 30(11), 1197-1217.

Yewlett, C. J. L. (2017). Operations research models for public administration. *Journal of the Operational Research Society*, 28(2), 354–355.

Zeller, C. (2004). North Atlantic innovative relations of Swiss pharmaceuticals and the proximities with regional biotech areas. *Economic Geography*, 80(1), 83-111.

Zhao, S. (2003). Toward a taxonomy of copresence. *Presence: Teleoperators & Virtual Environments*, 12(5), 445-455.

Zimmermann, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological development, and future prospects. *American Educational Research Journal*, 45(1), 166-183.

Zukin, S., & DiMaggio, P. E. (1990). *Structures of capital: The social organization of the economy*: CUP Archive.

APPENDIX 1.

Institutional Acronyms and Type

Name of Agencies	Denotation	Institution
Department of Agriculture	USDA	Executive Department
Department of Commerce	DOC	Executive Department
Department of Defense	DOD	Executive Department
Department of Education	ED	Executive Department
Department of Energy	DOE	Executive Department
Department of Health and Human Services	HHS	Executive Department
Department of Homeland Security	DHS	Executive Department
Department of Housing and Urban Development	HUD	Executive Department
Department of the State	DOS	Executive Department
Department of Justice	DOJ	Executive Department
Department of Labor	DOL	Executive Department
Department of Interior	DOI	Executive Department
Department of Transportation	DOT	Executive Department
Department of the Treasury	USDT	Executive Department
Department of Veterans Affairs	DOV	Executive Department
Environmental Protection Agency	EPA	Independent Agency
General Services Administration	GSA	Independent Agency
The United States Agency for International Development	USAID	Independent Agency
National Aeronautics and Space Administration	NASA	Independent Agency
National Science Foundation	NSF	Independent Agency
Office of Personnel Management	OPM	Independent Agency
Small Business Administration	SBA	Independent Agency
Social Security Administration	SSA	Independent Agency
Nuclear Regulatory Commission	NRC	Independent Agency

APPENDIX 2.

Examples of Archival Analysis: Efficiency

Value Content	Level	Agency (Year)	Examples
Efficiency	Strategic Goals	U.S. General Service Administration (2017)	Improve the efficiency of operations and service delivery.
		Department of Justice (2013)	Ensure and support the fair, impartial, efficient, and transparent administration of justice at the federal, state, local, tribal, and international levels.
	Strategic Objectives	Department of Defense (2017)	Improve overall performance, strengthen business operations, and achieve efficiencies, effectiveness, and cost savings that can be transferred to higher priority needs.
		Department of Transportation (2013)	Reduce carbon emission, improve energy efficiency, and reduce dependence on oil.
	Performance Indicators	United States Agency International Development (2017)	Number of participants in the young African leader's initiative.
		U.S. Department of Health and Human Services (2013)	Amount of savings by State ADAPs participation in cost-savings strategies on medications.

Examples of Archival Analysis: Service Quality

Value Content	Level	Agency (Year)	Examples
Service Quality-Reliability	Strategic Goals	Environmental Protection Agency (2017)	Protecting human health and the environment by enforcing laws and assuring compliance.
		Office of Personnel Management (2013)	Provide the training, benefits, and work-life balance necessary for federal employees to succeed, prosper, and advance in their careers.
	Strategic Objectives	Department of Education (2017)	Support implementation of internationally benchmarked college- and career-ready standards, with aligned, valid, and reliable assessments.
		Social Security Administration (2013)	Maintain secure and reliable IT services.
	Performance Indicators	Environmental Protection Agency (2017)	Improve and/or restore water and habitat quality to meet water quality standards in watersheds throughout the five Gulf States and the Mississippi River Basin.
		Department of Defense (2013)	Percent of DoD's nuclear command, control, and communications (NC2) cryptographic modernization plan completed.
Service Quality-Responsiveness	Strategic Goals	U.S. Office of Personnel Management (2017)	Provide timely, accurate, and responsive service that addresses the divers needs of our customers.
		Department of Health & Human Service (2013)	Protect Americans' health and safety during emergencies, and foster resilience in response to emergencies.
	Strategic Objectives	Environmental Protection Agency (2017)	Prepare for and respond to accidental or intentional releases of contaminants and clean up and restore polluted sites for reuse.
		Department of Defense (2013)	Improve the responsiveness and flexibility of consequence management response forces.
	Performance Indicators	U.S. Office of Personnel Management (2017)	Percent of external-facing program offices measuring their customer service timeliness.
		Social Security Administration (2013)	Minimize average wait time from hearing request to decision.

Examples of Archival Analysis: Customer Satisfaction

Value Content	**Level**	**Agency (Year)**	**Examples**
Customer Satisfaction	Strategic Goals	U.S. General Service Administration (2017)	Deliver excellent customer service.
	Strategic Objectives	Department of Treasury (2017)	Create a culture of service through relentless pursuit of customer value.
		Social Security Administration (2013)	Increase public satisfaction with our telephone services.
	Performance Indicators	Department of Veterans Affairs (2017)	Veterans' satisfaction level with the VA Loan Guaranty program (out of 1000).
		Office of Personnel Management (2013)	Percent of customers satisfied with overall retirement services.

Examples of Archival Analysis: Social Equity

Value Content	Level	Agency (Year)	Examples
Vertical Equity	Strategic Goals	U.S. Department of Interior (2017)	Improve the quality of life in tribal and native communities.
		Social Security Agency (2013)	Deliver quality disability decisions and services.
	Strategic Objectives	Department of Education (2017)	Equitable educational opportunities increase all students' access to educational opportunities, and remove barriers that students face based on their race, ethnicity, or national origin; sex; gender identity or expression; disability
		Small Business Administration (2013)	Strengthen outreach to underserved communities and populations.
	Performance Indicators	U.S. Agency International Development (2017)	Number of participants in the young African leader's initiative.
		Department of Transportation (2013)	Improved access to transportation for people with disabilities and order adults.
Horizontal Equity	Strategic Goals	Department of Health & Human Services (2017)	Make coverage more secure for those who have insurance and extend affordable coverage to the uninsured.
		U.S. Department of Agriculture (2013)	Ensure that all of America's children have access to safe, nutritious, and balanced meals.
	Strategic Objectives	Department of Treasury (2017)	Promote savings and increased access to credit and affordable housing options.
		Office of Personnel Management (2013)	Promote diversity and inclusion in the federal workforce.
	Performance Indicators	U.S. Agency International Development (2017)	number of people gaining access to basic sanitation services
		Department Education (2013)	Ensure equal access to education and promote educational excellence throughout the nation through the vigorous enforcement of civil rights laws.

APPENDIX 3.

Value Content Expressed in the PARs (2013)

	Value Content				Total
	Efficiency	Service Quality	Customer service satisfaction	Social Equity	
USDA	0 (0.0%)	14 (25.45%)	0 (0.0%)	11 (20.00%)	25/55
DOC	3 (3.66%)	21 (25.61%)	0 (0.0%)	5 (6.10%)	29/82
DOD	5 (4.95%)	32 (31.68%)	0 (0.0%)	1 (0.99%)	38/101
ED	0 (0.0%)	30 (49.18%)	1 (1.64%)	15(24.59%)	46/61
DOE	5 (14.29%)	17 (48.57%)	0 (0.0%)	0 (0.0%)	22/35
HHS	3 (1.81%)	49 (29.52%)	4 (2.41%)	51 (30.72%)	107/166
DHS	0 (0.0%)	26 (26.00%)	6 (6.00%)	7 (7.00%)	39/100
HUD	2 (1.98%)	19 (18.81%)	5 (4.95%)	31(30.69%)	57/101
DOS	7 (2.00%)	119 (34.00%)	4 (1.14%)	31 (8.86%)	161/350
DOJ	2 (7.69%)	3 (11.54%)	0 (0.0%)	3 (11.54%)	8/26
DOL	0 (0.0%)	32 (26.45%)	3 (2.48%)	11 (9.09%)	46/121
DOI	1 (1.33%)	13 (17.33%)	0 (0.0%)	7 (9.33%)	21/75
DOT	4 (4.30%)	10 (10.75%)	0 (0.0%)	11 (11.83%)	25/93
USDT	2 (3.77%)	15 (28.30%)	0 (0.0%)	3 (5.66%)	20/53
DOV	1 (2.63%)	12 (31.58%)	4 (10.53%)	0 (0.0%)	17/38
EPA	0 (0.0%)	28 (26.17%)	0 (0.0%)	17 (15.89%)	45/107
GSA	5 (16.13%)	2 (6.45%)	0 (0.0%)	1 (3.23%)	8/31
NASA	4 (2.25%)	86 (48.31%)	1 (0.56%)	8 (4.49%)	99/178
NSF	0 (0.0%)	12 (34.29%)	0 (0.0%)	0 (0.0%)	12/35
OPM	0 (0.0%)	12 (28.57%)	2 (4.76%)	8 (19.05%)	22/42
SBA	3 (6.00%)	12 (24.00%)	4 (8.00%)	8 (16.00%)	27/50
SSA	2 (3.77%)	19 (35.85%)	2 (3.77%)	19 (35.85%)	42/53
NRC	2 (3.03%)	41 (62.12%)	0 (0.0%)	0 (0.0%)	43/66
Total *N*	51 (2.53%)	624 (30.91%)	36 (1.78%)	248 (12.28%)	959/2019

Value Content Expressed in the PARs (2017)

	Value Content				Total
	Efficiency	Service Quality	Customer service satisfaction	Social Equity	
USDA	3 (4.92%)	10 (16.39%)	0 (0.0%)	14 (22.95%)	27/61
DOC	1 (1.10%)	15 (16.48%)	5 (5.49%)	4 (4.40%)	25/91
DOD	4 (5.13%)	27 (34.62%)	1 (1.28%)	2 (2.56%)	34/78
ED	0 (0.0%)	27 (31.76%)	0 (0.0%)	19 (22.35%)	46/85
DOE	11 (15.49%)	22 (30.99%)	1 (1.41%)	0 (0.0%)	34/71
HHS	3 (1.86%)	38 (23.60%)	4 (2.48%)	46 (28.57%)	91/161
DHS	0 (0.0%)	27 (21.95%)	7 (5.69%)	2 (1.63%)	36/123
HUD	2 (3.39%)	7 (11.86%)	0 (0.0%)	16 (27.12%)	25/59
DOS	4 (2.44%)	54 (32.93%)	2 (1.22%)	19 (11.59%)	79/164
DOJ	2 (3.92%)	10 (19.61%)	0 (0.0%)	8 (15.69%)	20/51
DOL	0 (0.0%)	35 (24.14%)	2 (1.38%)	10 (6.90%)	47/145
DOI	2 (1.89%)	14 (13.21%)	0 (0.0%)	21 (19.81%)	37/106
DOT	7 (7.87%)	18 (20.22%)	0 (0.0%)	6 (6.74%)	31/89
USDT	5 (7.81%)	19 (29.69%)	2 (3.13%)	6 (9.38%)	32/64
DOV	0 (0.0%)	10 (12.20%)	16 (19.51%)	10 (12.20%)	36/82
EPA	1 (0.56%)	46 (25.99%)	0 (0.0%)	17 (9.60%)	64/177
GSA	4 (11.11%)	3 (8.33%)	5 (13.89%)	5 (13.89%)	17/36
NASA	0 (0.0%)	9 (20.45%)	0 (0.0%)	0 (0.0%)	9/44
NSF	0 (0.0%)	11 (39.29%)	0 (0.0%)	3 (10.71%)	14/28
OPM	2 (1.57%)	38 (29.92%)	6 (4.72%)	7 (5.51%)	53/127
SBA	1 (3.45%)	6 (20.69%)	2 (6.90%)	5 (17.24%)	14/29
SSA	2 (3.64%)	40 (72.73%)	2 (3.64%)	15 (27.27%)	59/55
NRC	2 (4.00%)	29 (58.00%)	0 (0.0%)	0 (0.0%)	31/50
Total *N*	56 (2.83%)	515 (26.06%)	55 (2.78%)	235 (11.89%)	861/1976