

THREE ESSAYS ON BUREAUCRATIC REPUTATION:
PREDICTORS, MEASURES, AND STRATEGIES

By

DANBEE LEE

A Dissertation submitted to the
Graduate School-Newark
Rutgers, The State University of New Jersey

In partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
Graduate Program in Public Affairs and Administration

Written under the direction of

Gregg G. Van Ryzin

And approved by

Newark, New Jersey

May 2020

Copyright page:

©[2020]

Danbee Lee

ALL RIGHTS RESERVED

ABSTRACT OF THE DISSERTATION PROPOSAL

Bureaucratic reputation in the eyes of citizens:

Predictors, measures, and strategies

by DANBEE LEE

Dissertation Director:

Gregg G. Van Ryzin

Bureaucratic reputation has been defined as a set of beliefs about an organization's capacities, intentions, history, and mission that are embedded in a network of multiple audiences (Carpenter, 2010, 45). An agency's reputation is closely linked to its level of approval and support, allowing the agency to establish autonomy, accumulate power, and enhance legitimacy when it is effectively managed (Maor, 2016). Also, previous studies have shown that the reputational concerns of agencies shape their behaviors, such as accountability, collaboration, and communication strategies (Busuioc, 2016; Gilad, Maor, & Ben-Nun Bloom, 2013; Ingold & Leifeld, 2014).

While scholars have paid increasing attention to the consequences of bureaucratic reputation for the behavior and autonomy of public organizations, little is known about: 1) measuring bureaucratic reputation in the eyes of citizens and other audiences, 2) predictors of citizens' reputation judgments of various agencies, and 3) strategies to shape audiences' reputation judgments. All of these issues are key to understanding and thus managing bureaucratic reputation. Given that knowing how reputations are formed and cultivated

remains “fundamental to understanding the role of public administration in a democracy” (Carpenter & Krause, 2012, 26), more research on bureaucratic reputation from the viewpoint of citizens is needed.

This dissertation probes three different research questions related to the construction of bureaucratic reputation in the eyes of citizens. The theoretical frameworks are grounded upon both the accountability and management literatures, particularly governance theory and New Public Management. Based on these perspectives, the first essay focuses on developing and validating reputation measurements through an empirically grounded scale development process. The second essay explores predictors of citizens’ reputation judgments of U.S. federal agencies using national survey data. The third essay examines reputation management, particularly the effect of communication strategies on citizens’ reputation judgments, utilizing a set of survey experiments.

The results provide empirically grounded tools and ideas that help better understand bureaucratic reputation at the individual level of the citizenry, perhaps the most important audience of any bureaucracy. First, the results provide a validated scale to measure bureaucratic reputation that can be used in surveys and other studies. Second, the results show how citizens’ characteristics shape their reputation judgments of federal agencies. Third, the results demonstrate how citizens’ reputation judgments can be managed through agency communication strategies after a policy or administrative failure. Taken together, these findings enhance the theoretical understanding of bureaucratic reputation and provide public agencies with empirical implications for managing their relations with the public.

ACKNOWLEDGMENT

My dissertation could never have been written without the help of my professors, colleagues, family members, and friends. I would like to express my gratitude to them. They were there since I began to develop research ideas, providing comments and consult on the revised models. They also enriched my academic life by providing wonderful opportunities, including attending conferences, linking me to academic networks, and advising me on the essentials of publishing papers in academic journals. I would thank everyone who has been there for me whenever I have been in need.

First, I would like to thank my adviser, Dr. Gregg G. Van Ryzin for being a great academic facilitator who allowed me to transform my research interests in organizational reputation into academic work. Thank you for your encouragement and intellectual guidance whenever I had concerns related to my continued research on this topic. You have been a wonderful academic mentor who has encouraged me to overcome any difficulties I encountered. It has been a pleasure working with you and an honor to receive your wonderful advice.

I was fortunate to meet my PhD committee member, Dr. Sebastian Jilke at Rutgers, who has provided enthusiastic support of my research and academic life. You have been like an older brother during my time here. I will miss your straightforward communication and strong encouragement. You provided invaluable help to break out of my shell with everything in my academic life, including research, networking, and finding my way on the job market. It has been fun talking with you, and I get energy and inspiration from all of our conversations.

I could not have started this dissertation without meeting Dr. Frank J. Thompson in 2015 when I took his class and encountered the topic of organizational reputation. Thanks to your encouragement, I developed my interest in this area into this dissertation. I always felt empowered and full of confidence after talking with you. I will always remember where my research initiated. Thank you for being my PhD committee member and your ceaseless support of my pursuit of research on this topic.

I would also like to extend my gratitude to Dr. Sharon Gilad, my external committee member. I am grateful that you agreed to be part of my PhD committee, and thank you for providing me with wonderful opportunities. I had unforgettable moments at ECPR in 2018 attending the reputation panel you co-organized that enabled me to meet with great reputation researchers. In addition, I really appreciate that you gave me an opportunity to be a referee, extended your time to me at PMRC, and gave invaluable feedback on my studies. I am also thankful for your instrumental advice in your area of expertise, the topic of reputation. I learned a lot from you.

I would also like to thank professors at Rutgers SPAA, Drs. Yahong Zhang, Jiahuan Lu, Gregory Porumbescu, Linsey McDougale, Alan Sadovnik, Pengju Zhang, Cleopatra Charles, , Jongmin Shon, who I had a lot to learn by working with them. Thanks for being very supportive and giving me advice in every essential step of this process. My thanks also go out to my SPAA colleagues, Chengxin Xu and Ivan Lee, I really enjoyed talking with them and have learned a lot from them, as well. I would also like to thank Seulki Lee, Sinah Kang, and Cheon Lee and appreciate everyone else at SPAA who made my time enjoyable and meaningful.

My heartfelt thanks go to my lifetime partner, my husband Minsung Kang for all the help, support, and love that have made me who I am today. My PhD journey over the past five years would have been impossible without you being there for me. Thank you so much for your unconditional support and all the intellectual advice you have given me. I say that it is not “my” effort but “our” effort that made this achievement possible. I would specially like to thank you for your jokes that made me laugh every day. My life would be meaningless if you were not there.

Special thanks go to my parents, brother, and friends. I would like to thank my mother and father for their encouragement and for being there every time I need them. They are my role models, and I cannot thank them enough for their unconditional support of my life, work, and future. I would also like to express my thanks to my mother-in-law, father-in-law, and brother-in-law for being such big supporters in my life. In addition, I would like to thank my friends for understanding my ups and downs with care. Their great advice and emotional support are the best remedies for hard times.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION: UNDERSTANDING BUREAUCRATIC REPUTATION.....	1
1. Introduction.....	1
1.1 Definitions of bureaucratic reputation.....	2
1.2 Importance of bureaucratic reputation.....	5
2. Bureaucratic Reputation Literature.....	5
2.1 Reputation and accountability.....	5
2.1.1 Reputation in administration-politics relationships.....	6
2.1.2 Reputation in governance.....	9
2.1.2.1 Governance as networks.....	10
2.1.2.2 Governance as New Public Management.....	13
2.2. Reputation and management.....	14
2.2.1 Resource-based View.....	15
2.2.2 Reputation management: Selective responses.....	17
2.2.3 Reputation management: Selective communication strategies.....	19
2.3 Methodological approaches.....	20
3. Roadmap of the Dissertation.....	23
3.1 Gaps in the bureaucratic reputation literature.....	23
3.2 Research purpose.....	26
3.3 Research questions.....	28

CHAPTER 2

MEASURING BUREAUCRATIC REPUTATION.....	32
1. Summary.....	32
2. Introduction.....	32
3. Research Background.....	33
3.1 The concept of bureaucratic reputation	33
3.2 The measurement of reputation	37
4. Scale Development Processes.....	39
4.1 Item generation.....	39
4.2 Expert review.....	40
4.3 Citizen survey.....	40
4.4 Factor analysis and reliability test.....	42
4.5 Construct validity.....	48
5. Agency Levels and Profiles.....	52
6. Conclusions	56

CHAPTER 3

PREDICTORS OF CITIZENS' REPUTATION JUDGMENTS	58
1. Summary	58
2. Introduction	59
3. Research Background	60
4. Data and Measurement	67
4.1 Outcome variable	67
4.2 Predictor variables	69
5. Analysis and Results	74
6. Conclusions	77

CHAPTER 4

REPUTATION MANAGEMENT AND COMMUNICATION STRATEGIES	80
1. Summary	80
2. Introduction	80
3. Research Background	83
3.1 Blame avoidance strategies and bureaucratic reputation	83
3.2 Hypotheses and models	86
4. Research Design	91
5. Analysis and Results	97
6. Conclusions	105

CHAPTER 5

DISCUSSION, LIMITATIONS, AND IMPLICATIONS	109
1. Introduction	109
2. Summary of Findings	110
2.1 Measuring bureaucratic reputation: Scale development and validation	111
2.2 Formation of citizens' reputation judgments	112
2.3 Reputation management and communication strategies	113
3. Limitations	114
3.1 Methodological limitations	114
3.2 Theoretical limitations	116
4. Implications for Future Agenda	119
4.1 Implications for theory and research	119
4.2 Implications for policy and practice	125
4.2.1 The way to maintain broad-based support for an agency	126
4.2.2 The way to steer a vessel amid hazardous shoals	127
4.2.2.1 Conflict of audiences' interest	127
4.2.2.2 Coping with failures	128
5. Closing remarks	129

REFERENCES.....	131
APPENDIX A.....	144
Description of the sample. Chapter 2	
APPENDIX B.....	145
Exploratory Factor Analyses of Split-Half Samples. Chapter 2	
APPENDIX C.....	146
Bureaucratic Reputation Scale. Chapter 2	

LIST OF TABLES

Table 2.1: Exploratory Factor Analysis.....	44
Table 2.2: Correlation analysis.....	49
Table 3.1: Question wording and descriptive statistics for independent variables.....	71
Table 3.2: Logit regression analyses of federal agency reputations.....	73
Table 4.1: Classification of failure types and matched response.....	89
Table 4.2: Final sample descriptive statistics.....	96
Table 4.3: T-test of manipulation check of the scenario.....	97
Table 4.4: Regression analyses of the effect of response strategies on bureaucratic reputation.....	103

LIST OF FIGURES

Figure 2.1: Confirmatory Factor Analysis for 5-item Scale.....	46
Figure 2.2: Confirmatory Factor Analysis for 10-item Scale	47
Figure 2.3: 5-items Reputation Scale Score.....	53
Figure 2.4: Dimensional Profile of Reputation Measure.....	55
Figure 3.1: Conceptual model of citizen characteristics as determinants of agency reputations.....	63
Figure 3.2: Reputations (favorability ratings) of 12 federal agencies (in rank order).....	69
Figure 4.1: An experimental flow chart.....	94
Figure 4.2: Mean comparison of pre and post reputation score by failure types.....	99
Figure 4.3: Mean comparison of pre and post reputation scores by strategic responses that are sub-grouped by failure types.....	101

CHAPTER 1. INTRODUCTION: UNDERSTANDING BUREAUCRATIC REPUTATION

1. Introduction

1.1 Definitions of bureaucratic reputation

There are various definitions of *organizational reputation*, including those specific to economics, business, and sociology (Chun, 2005; Fombrun, Gardberg, & Sever, 2000). According to Walker (2010), the definition referenced most frequently is that of Fombrun (1996) found in the business literature, where he defines “organizational reputation” as the “...collective representation of a firm’s past actions and results that describe the firm’s ability to deliver valuable outcomes to multiple stakeholders” (p. 72). In 2012, Fombrun redefined this definition as “A corporate reputation is a collective assessment of a company’s attractiveness to a specific group of stakeholders relative to a reference group of stakeholders relative to a reference group of companies with which the company competes for resources” (p. 100).

However, this dissertation relies principally on Carpenter’s definition (2010), which focuses on public rather than private organizations. Carpenter defines *bureaucratic reputation* as “...a set of symbolic beliefs about the unique or separable capacities, roles, and obligations of an organization, where these beliefs are embedded in audience networks” (p. 45). While the definitions of both public and private organizations’ reputations stress their collective judgment, the business and economics literatures focus more on performance as organizations’ economic gains. In contrast, bureaucratic reputation accentuates performance with respect to values that are considered important in a

democratic society—morality, fairness, and justice, among others. This dissertation follows Carpenter's (2010) work and bases its research questions and models on his approach.

Besides bureaucratic reputation's definitions, the discussion of reputation as a construct should be elaborated. Although Carpenter (2010) defined reputations as 'beliefs,' which is a more of a sticky construct, still reputation studies sometimes consider it in a more flexible way because there has been no one sources that captures the entire concept of reputation, and definitional problems remain in the literature (Chun, 2005). Among many different definitions, this study views reputation as at least partly attitude—a person's feeling toward and evaluation of some object—rather than a belief—a person's subjective probability that an object has a particular attribute (Fishbein & Ajzen, 1975). One of the reasons is that this study uses an approach that views reputation is a combination of affect and evaluation of the organization, where it is more changeable depending on external changes, which are attributes of attitude. Further, this study views opinions and attitudes in an identical way, following McGuire (1985). Unlike beliefs, which describe people's reaction as true/false toward the information, according to Fishbein & Ajzen (1975), reputation is the reaction of agree or disagree that depends on individuals' affect and evaluation of organizations.

1.2 Importance of bureaucratic reputation

Why does bureaucratic reputation matter in public administration? There are two approaches to answer this question. The first is to view reputation as part of an *accountability mechanism*, particularly as external sanctions on public organizations'

behaviors. When referring to accountability, this study follows Bovens' (2007) definition and confines the discussion to the role of reputation in the relationship between "actors" and "forums." Bureaucratic reputation scholars follow this approach, with an emphasis on administrative-politics relationships. In this approach, when a public agency builds a good reputation with external audiences, such as politicians, the agency is trusted to behave appropriately and is given more power than control. Carpenter (2001) conducted a groundbreaking study based on this approach, and subsequent empirical evidence has endorsed the effects of reputations on administrative agencies' deviation from traditional political control.

In addition to administrative-politics relationships, reputation-based accountability mechanisms can be applied to governance theory also, particularly in the form of networks and New Public Management (NPM). Following Rhodes (1996), who defined governance as self-organizing and inter-organizing networks, and a third governing structure in addition to market and hierarchy, network relationships are characterized by trust, reciprocity, and mutual interdependence (Larson, 1992; Thompson, Frances, Levacic, & Mitchell, 1991). Based on this approach, reputations are considered one of the central facilitators of coordination among participants in network relationships.

In addition, with respect to governance as NPM, which emphasizes a performance-oriented regime, government outcomes are a critical aspect of accountability. NPM also stresses citizens' role as an external controller of administrative agencies that is similar to elected officials who pass judgments on, and pose questions about, actors' conduct, in this case, governments (Bovens, 2007; Pollit, 2011; Osborne & Gabler, 1993; Van de Walla, 2016). According to this approach, reputations can be indicators of public organizations'

performance that are similar to citizens' satisfaction or trust in government, which will be discussed further in the next section.

The second justification for studying bureaucratic reputation is managerial. This approach views organizational reputation as one of an organizations' invisible assets that is a key part of environmental information—the information flow from the environment to organizations (Itami, 1987). Further, Carpenter (2010) described reputations as political assets that protect organizations from political attacks, generate public support, achieve delegated autonomy and discretion from politicians, and help recruit and retain valued employees. While accountability-based and management-based approaches overlap in certain ways, this research distinguishes the two: the accountability-based approach focuses on reputation's role in relationships between “actors” and “forums” while the management approach investigates reputation's effect on organizations, as well as the way organizations manage their reputations as either a resource or environment.

Specifically, this research confines the discussion of reputation management to three streams. The first is grounded in the Resource-Based View that emphasizes the role of a resource as a source of comparative advantage and explains the effect of reputations on organizations' effectiveness, including performance (Barney, 1991). The second and third streams focus on managing organizations' relationships with audiences in their environment. This approach is based on the view that reputations are sources that are drawn from organizations' environments, particularly in the interactions with their audiences (Carpenter, 2010; Fombrun, 1990). Following Maor's (2015) categorization, this research distinguishes the second and third streams. While the principal idea of both is managing reputation, the second focuses on organizations' selective responses, while the third centers

on communication strategies. Briefly, selective responses are organizations' behavioral changes, while communication strategies are their use of argument in response to their reputational concerns.

The following section includes three subsections that discuss the theoretical and methodological approaches in the bureaucratic reputation literature in-depth, and the third subsection is composed of two further subsections. The next section, The Roadmap of this Dissertation, includes three subsections. The first discusses gaps in the current reputation literature and the theoretical and methodological approaches this dissertation used to fill those gaps. In its theoretical approach, this study followed the accountability and reputation management approaches grounded in NPM, which considers citizens key audiences in formulating reputations. Methodologically, this research adopted the individual-level approach using quantitative methods, including scale development, an observational study, and a survey experiment. The research purpose and questions that are based on the theoretical and methodological approaches are discussed subsequently, as well as the way those contribute to the public administration literature.

2. Bureaucratic Reputation Literature

2.1 Reputation and accountability

Among many different accountability definitions in the literature, this study followed the approach that considers accountability "...a relationship between an actor and a forum in which the actor is obliged to explain and justify his conduct; the forum can pose questions; pass judgment; and the actor may face consequences" (Bovens, 2007, p. 452). Based on this approach, according to Romzek and Dubnick's (1987)

accountability typology, bureaucratic reputation can be a source of political accountability that includes political, bureaucratic, legal, and professional accountability, as they stated that a source of political accountability represents the democratic pressure the general public, elected officials, and interest groups exert. Because bureaucratic reputation includes evaluations overall on the part of multiple stakeholders, a function as an external sanction on bureaucratic behaviors, it can also be categorized as a political mechanism.

2.1.1 Reputations in administration-politics relationships

The Fredrich-Finer debate represents two contrasting views of bureaucrats' accountability that remains unresolved. Fredrich (1940) emphasizes public servants' accountability to professionalism and public sentiment, while Finer (1941) focuses on bureaucrats' accountability to elected officials. This debate relates more to the process by which public administration copes with environmental changes. In traditional bureaucratic accountability in a representative democracy, administrative agencies' behaviors are under the control of elected officials, while contemporary government faces increasing demands to enlarge its service areas, solve difficult social problems, and manage multiple stakeholders who participate in decision making, public service delivery, and implementation (Fredrich, 1940). Fredrich championed public administrations' novel accountability in contemporary government, while Finer upheld their traditional accountability.

A group of bureaucratic reputation scholars has focused on *administration-politics relationships*, particularly what reputations can do to improve administrative

autonomy. There are contradicting views on whether bureaucratic autonomy increases agencies' capacity to make welfare-improving decisions or to act without regard to the public's preferences (Miller and Whiteford, 2016). Despite the contradiction, both views agree that the autonomy can be an incentive or sanction on bureaucratic behavior, as the discussion of the role of bureaucratic reputation has evolved in the Finer-Friedrich debate (accountability vs. professionalism). Specifically, the assumption under this approach is that reputations can be informal tools to control bureaucrats and, even further, for bureaucrats to earn autonomy coming from one of its key sources—public support. This approach deviates from the traditional principal-agent relationship, which emphasizes elected officials' direct control over administrations based on formal rules.

The early scholarly views on bureaucratic reputation were consistent with “control.” Moe (1984) suggested that bureaucrats' reputations can be “...a central tool in the task of political monitoring” (p. 767) that can serve as alternatives for politicians to control bureaus when there is information asymmetry within principal-agent relationships. Consistent with the conventional approach to administration-politics relationships, which focuses primarily on the way elected officials control public agencies' behaviors, bureaucratic reputation has been suggested to “facilitate the monitoring” of bureaucrats without formal control because it gives public agencies an incentive to have more “...delegated autonomy and discretion from politicians” (Carpenter, 2002, p. 491; Moe, 1984, p. 767). In their views, when a public agency builds a good reputation in its relationships with external audiences, such as politicians, the agency is trusted to behave properly and is given more power than control, which leads ultimately to accountable behaviors.

Nevertheless, later work has provided evidence that an organization's reputation becomes a source of bureaucratic autonomy that assigns more power to public organizations in their relationships with elected representatives and their control. Carpenter's (2001) groundbreaking work on bureaucratic reputation, and his subsequent research has found empirical evidence of reputation's effects on administrative agencies that deviate from traditional political control. The author stressed that bureaucratic reputation made the Department of Agriculture and the Post Office Department successful, while the Department of the Interior failed to achieve autonomy. Indeed, Krause and Douglas (2005) showed that bureaucratic reputation is a stronger driver of agency production than other structural designs insulated by political control. Similarly, Maor (2007) suggested that drug reimbursement organizations used scientific 'gold' standards to protect their reputation from the public than from political authorities. These studies all concur that bureaucratic reputation motivates bureaucrats to behave in ways that deviate from traditional administration-political relationships.

While some scholars have suggested the possibility that reputation is a source of bureaucratic autonomy, there are counterarguments as well. Some view bureaucratic reputation as an unstable source to secure bureaucratic autonomy (Moar, 2015). Roberts (2006) asserted that bureaucratic reputation is vulnerable to politicization based on the case of FEMA, which enjoyed only short-term autonomy attributable to its positive reputation. Maor (2015) and Wæraas and Byrkjeflot (2012) made a further criticism of reputation's role in securing autonomy: a public agency's good reputation is insecure because of the institutionalized pressure of political control. For example, if political authorities decide to change their priorities, public agencies are vulnerable to political interference. While the

long-term effect of bureaucratic reputation on agencies' autonomy is still disputable, many scholars agree that its effect is temporary.

Beyond these discussions of the way reputations cultivate bureaucratic autonomy and influence public administration's political accountability to elected representatives, some researchers focus to a greater degree on explaining the *reputation-accountability mechanism* in a broader setting that involves "audiences." Busuioc and Lodge (2016) suggested that reputation is a critical driver of bureaucrats' accountable behavior, as it leads agencies to act upon the reputational calculation to alleviate external influences, such as political power and control. They explained that the existence of "multiple audiences" shifts traditional accountability that focuses on one-sided relationships between forums and actors to accountability in "...advancing one's reputation vis-à-vis different audiences" (Busuioc & Lodge, 2016, p. 92). Thus, the process of managing reputation involves forums and actors who prioritize each other among multiple audiences, and thus shape bureaucratic accountability in a more interactive way.

2.1.2 Reputation in governance

Government refers to "...the formal institutional processes which operate at the level of the nation-state to maintain public order and facilitate collective action" (Stoker, 1988, p. 17). On the other hand, governance signifies "...a change in the meaning of government, referring to a new process of governing; or changed condition of the ordered rule; or the new method by which society is governed" (Rhodes, 1996, p. 652). The output of governance is similar to that of government, which is "...the conditions for ordered rules and collective action," but the differences lie in its process (Stoker 1998, p. 17). Unlike the

government, which centers on authority and a sanction-centered governing style, Stroke (1998) suggested that the baseline agreement of governance is “...the development of governing styles in which boundaries between and within the public and private sectors have become blurred” (p. 17). Further, Rhodes’ (1996) definition of governance is a “...self-organizing and inter-organizational network,” and shared characteristics of governance as followings: 1) interdependence between organizations involving public, private and nonprofit sectors 2) the need to exchange resource drove the continuing interactions between network members 3) game-like interactions rooted in trust while the rule of the game is drawn upon negotiation and agreement by network participants; autonomy from the state.

The current public administration literature includes numerous definitions of governance; however, as discussing the definition of governance is not the main topic of this research, this study adopts Frederickson’s (2005) broad approach that describes governance as “...most descriptions of governance—networks, inter-organizational and inter-jurisdictional cooperation, power-sharing federation, public-private partnerships, and contracting out—are forms of institutional adaptation in the face of increasing interdependence” (p. 290). For theoretical clarification, this study addresses the views of governance as Networks and NPM, where both allow the reputation discussion to fit within the context of governance.

2.1.2.1 Governance as networks

The network governance approach views governance as a broader concept than government, as it emphasizes service delivery in the form of a network, an inter-

organizational association that involves a complex set of private and public organizations that need to exchange resources to achieve their objectives (Rhodes, 1997; Rosenau, 1992). In the public administration context, networks refer to "...bureaucratic structures of interdependence involving multiple organizations where institutional glue congealing networked ties may include authority bonds, exchange relationships, and common-interest based coalitions" (O'Toole, 1997, p. 445).

One of the important factors in this network form of governance is reputation, in addition to trust, reciprocity, and mutual interdependence (Larson, 1992). According to Thompson, Frances, Levacic, and Mitchell (1991), "If it is price competition that is the central coordinating mechanism of the market and administrative orders that of hierarchy, then it is trust and cooperation that centrally articulates networks" (p. 265-76). In the network setting, where reciprocity is important in interdependent organizations' relationships, a reputation for being trustworthy can be a strong driver that maintains the participants' accountable behavior as an important asset and a mutually reinforcing core (Ostrom, 1998).

Consistent with the approach that emphasizes coordination in public service delivery, Van Slyke (2007) conducted an empirical study that focused on contracting to investigate whether reputations are important internal rewards and effective sanctions that make nonprofit organizations more accountable in providing public services. Based on semi-structured interviews, he found that nonprofit managers, particularly in large nonprofit organizations, consider that enhancing their reputation is an important factor that brings benefits. With respect to public managers, he also confirmed that they use reputation as "decision heuristics" as a proxy for service quality and expected goal alignment (p. 176).

While this research focused on the role of reputations in nonprofit-public contracts, Busuioc's (2015) comparative study of two European Union agencies focused more on the role of reputation in inter-agency relationships. Using interviews, the study also provided empirical evidence that reputation calculations, which is cooperation's benefits to agency reputations' uniqueness, can be a factor that influences the success and failure of inter-agency cooperation.

Different from these approaches, which have applied reputation to coordination, other research has considered network governance from the view of blame avoidance, and emphasized factors that disrupt coordination. Networks include multiple interdependent actors with different preferences, which are characteristics that often lead to tension between protecting individuals' reputations over that of networks' or other network members (Moynihan, 2012). As a network blurs the responsibilities for outcomes and makes it difficult to identify who is accountable for results (Milward & Provan, 2000), it leads actors to use more blame avoidance strategies (Moynihan, 2012). The author suggested that an extra-network reputation exists, and argued that an agency weighs intra- and extra-network reputations differently, and that reputation-based power is often vulnerable to a change in extra-network reputation.

Moreover, there is another view that considers reputations as predictors of actors' power in networks, where the network system is characterized by ambiguous responsibilities and relies on the actors' perceived power, because its authoritative structure is more diffuse than that in a traditional hierarchy (Ingold & Leifeld, 2014). Thus, the perception of the power of political actors' influence differs from the true action of their political power. Ingold and Leifeld (2014) defined reputation as "...the perceived

importance of actors when evaluated by their peers or other stakeholders involved in the policy process” (p. 3) and analyzed the critical role of reputation comprehensively in both the vertical and horizontal integration of actors in the political system. By studying the factors that influence political actors’ reputation, they found that power and structural positions (informal and non-hierarchical powers) granted institutionally are predictors of political actors’ reputations.

2.1.2.2 Governance as New Public Management

The meaning of NPM can be characterized by two terms, managerialism and the new institutional economics (Hood, 2001; Pollit 2011; Rhodes, 1996). According to Osborne and Gaebler (1992), NPM is a transformation from “less government” to “more governance” that emphasizes competition, markets, customers, and outcomes. NPM’s proposed doctrines are: 1) managerialism represents adopting private sector management methods in the public sector, such as hands-on professional management, explicit standards and measures of performance, managing results, value for money, and closeness to customers, while 2) new institutional economics refers to incentive structures that include greater competition through contracting-out and quasi-markets, and customer choice (Hood, 1991; Rhodes, 1997). NPM suggests that management is appropriate for service delivery, while Osborne and Gaebler distinguish decision making and service delivery.

On the one hand, NPM overlaps with governance. According to Peters and Pierre (1998), both models share the following elements: 1) reduced role of elected officials; 2) current government is too distant from citizens and society, which leads to inefficient outcomes, and 3) the concept of steering, in which steering means setting broad policy

objectives, is important for public administration. On the other hand, they also stated the differences between NPM and governance: 1) governance is about the process, while NPM is about outcomes; 2) NPM is grounded in organizational and public choice theories, while governance is based on a theory of politics, and 3) governance does not share the core motivation to bring about a market-based revolution in the public sector, as does NPM, but to force the public sector to engage in a cooperative setting with the private sector. In addition to these differences between governance and NPM, Rhodes (1997) pointed out that the role of citizens in NPM occupies a stronger position, as customers can be direct evaluators of public agencies, and governance empowers citizens, but with limitations.

While this section described NPM in comparison to governance, the application of bureaucratic reputation to NPM will be discussed further in the next session that explains this dissertation studies' theoretical ground and contributions that are founded on NPM, particularly the role of the citizens.

2.2 Reputation and management

As reputations are influential resources of organizations, public organizations invest effort in building their reputations. When reputations are formed, cultivated, and managed successfully, they can become valuable political assets, as discussed in the previous section (Maor, 2015; Carpenter 2001). Because of its importance, various previous studies have focused on studying reputation management. Specifically, this study divides reputation management into two groups, resource management and environment management. Resource management applies the resource-related theory or framework,

while environment management focuses more on organizations' responses to reputational threats or concerns.

2.2.1. Resource-based View

The Resource-based View (RBV) is based on the assumption that organizations' comparative advantage is attributable to valuable, unique, rare, and sustainable resources that enhance the organization's performance ultimately on a short-term basis (Barney, 1991). This approach indicates that as organizations possess more resources, they enjoy greater growth and enhanced performance (Williamson, 1999). There are many different types of organizational resources, where "...resources include all assets, capabilities, organizational processes, firm attributes, information, knowledge, etc., controlled by a firm that enable the firm to conceive of and implement strategies that improve its efficiency and effectiveness" (Daft, 1987; cited in Barney, 1991, p. 101).

According to the RBV, an organizational resource has to meet certain criteria to be a source of organizations' sustained comparative advantage, indicating that an organization is "...implementing a value-creating strategy not simultaneously being implemented by any current or potential competitors" (Barney, 1991, p. 102). Based on the RBV, organizational reputation is an essential resource that is built upon internal investment and external appraisals (Dowling, 2001; Roberts & Dowling, 2002; Shamsie, 2003; cited in Boyd, Bergh, and Ketchen, 2010), which are also valuable, rare, and difficult to imitate (Boyd, Bergh, & Ketchen, 2010). Further, previous studies that have applied the RBV have provided evidence that organizational reputation can be a source of organizations'

competitive advantage that leads ultimately to superior performance (Hall, 1992; Roberts & Dowling, 2002).

In the public administration literature, Carpenter (2002) proposed that bureaucratic reputations "...are valuable political assets—they can be used to generate public support, to achieve delegated autonomy, and discretion from politicals, to protect the agency from political attack and to recruit and retain valued employees" (Carpenter, 2002, p. 491). Carpenter (2001) also argued that "...government agencies are motivated by a desire to demonstrate reputation uniqueness, which refers to their ability to create solutions and provide services found nowhere else in the polity" (p. 5). This argument indicates that reputations also bring government organizations unique characteristics that allow them to be differentiated from other similar ones, which is consistent with the RBV approach's assumptions about reputations.

Lee and Whiteford (2011) classified reputation as an important resource of public organizations, in addition to governments' five other types of resources, including administrative, human, financial, physical, and political. They stated that reputations are also a vital intangible resource of public organizations as "...those organizations compete in political reputation markets for the attention and support of the public and politicians" (p. 690). Using data from U.S. federal agencies, their empirical findings showed that reputations were one of the key resources that were related significantly to agencies' performance.

Despite the RBV's potential to enrich the public administration literature, similar to its influence on the business literature, current interests in the approach, particularly applying it to link reputations to performance, are rare in the field. In the condition under

which societal changes lead public administration to be responsive to different stakeholders, the RBV can help public organizations “...achieve a better fit or alignment with the demands and opportunities of their environments and ...perform better in those environments than organizations that do not” (Bryson, Ackermann, & Eden, 2007, p. 713). Thus, more studies are expected to apply the RBV in the bureaucratic reputation literature to identify ways to enhance organizations’ performance and further effectiveness.

2.2.2 Reputation management: Selective responses

According to Maor (2015), reputation management can be divided into two forms. The first is reputation management through changes in behaviors, including decision timing and observability, and agency outputs. The second is reputation management through the strategic use of communication. The first approach has been discussed actively in the public administration literature. Carpenter (2002) estimated the time that the U.S. Food and Drug Administration (FDA) took to approve a new drug, and although the review times for drugs differ, he found that political influence, particularly organized interest, caused considerable variation in review times. This also shows that the FDA can select audiences strategically to manage its reputation within those relationships. Like Carpenter (2002), Maor (2010) focused on the FDA’s timing of the jurisdictional claim when announcing its statutory authority in over-regulating novel technologies. The author found that the agency changed its actions if faced by reputational threats from other organizations and the public. Maor and Suilitzeanu-Kenan (2013) also found that reputational concerns affected the speed of the FDA’s response and that the valence of media coverage was associated with shorter response times.

Further, the reputation calculation influences public agencies' output directly. Maor and Sulitzeanu-Kenan (2016) defined responsiveness to change as "...greater volatility in agency activities when the organization is faced with a reputational threat and low volatility when reputation is not threatened" (p. 3), and their study showed that general outputs changed following an agency's reputational concerns, as in the case of Australia's federal welfare agency. In this case, the authors found that the agency tried to perform better to maintain its reputation even when it lacked political control. In addition, Rimkute's (2018) comparison of two cases of organizations charged with food safety, the European Food Safety Authority (EFSA) and the France food safety agency (ANSES), suggested that reputational threats influence public organizations' scientific assessment of chemicals as well.

Although many studies have concurred that the role of reputation is critical to improve the behavior of public agencies, as Moore (2015) showed, public agencies can nevertheless improve their performance and expand their programs even if their reputation is poor. Moore gave the example of the Veterans Affairs (VA), which endured several scandals and had a longstanding reputation for incompetence, but still achieved a major expansion of its healthcare program, and gained political support and increased funding. However, it is clear that overall, the studies cited above have shown that reputational concerns affect the behavior of public agencies, and often positively. Indeed, because agencies are sensitive to external attacks on their reputations, they try to secure their reputations by adjusting their performance. Thus, the risk of reputation-based punishment leads to greater accountability for their tasks and mission-oriented behaviors.

2.2.3 Reputation management: Selective communication strategies

Reputation management has both a ‘defensive’ component, in the sense of alleviating blame when reputation is threatened, and an ‘offensive component that tries to influence the content of public policy (Waeraas & Byrkjeflot, 2012; Christensen & Lodge, 2016). The current bureaucratic reputation literature focuses largely on defensive communication strategies that attempt to provide an effective response to reputational threats, restore public trust, and allow public agencies to acquire greater decision-making powers. Because of the public’s negative bias—citizens focus more on the government’s losses than gains—an agency needs to respond strategically when its reputation is at stake (Moynihan, 2012). This gives governments a strong incentive to avoid blame by using different communication strategies.

Researchers have attempted to show that public organizations “...carefully design their interaction with the public and shapes the ‘common ground’ that is shared with a critic” and “...it selectively responds in a way that retains its credibility” (Maor, 2015, p. 29). By analyzing the case of Hurricane Katrina, Moynihan (2012) suggested three strategies (blame-shifting, solo action, jurisdiction claiming)—all of which public agencies have used to protect their extra-network reputations. While Moynihan focused on extra-reputational threats, Maor et al. Bloom (2013) studied the use of different strategies and found that government agencies were more likely to remain silent if the criticism was about issues in which they had a strong reputation than if they had a weak reputation. The authors also presented different conditions with respect to various communication strategies, and explored carefully the contents of allegations in media coverage according to type of reputation (under- vs. over-regulation), salience, and severity of the issues, and the relative

importance of the reputation at stake. Their study expanded the discussion of communication strategies by describing more sophisticated types of silence, problem denial, full problem admission, and blame-shifting. All of these studies have helped enhance our understanding of the way public agencies use different strategic responses to safeguard or improve their reputations.

While most studies have focused on organizations' responses to a reputational threat, some have emphasized using communication strategies to enhance organizations' reputation. Christensen and Lodge (2016) investigated the symbols that public organizations use to manage their reputations by examining the kinds of information they post on their websites, believing that this shows the "soft power of the state" (p. 10). They found that public agencies manage different reputations strategically, such as by performance, moral and technical competence, or procedures. This has implications for the literature, insofar that the researchers analyzed the websites as a source of data that links the organization to the citizen directly, and the study showed that such activities can improve an agency's reputation, particularly in the public's eyes.

2.3 Methodological approaches

Most of the bureaucratic reputation studies have relied on qualitative studies, such as case studies, archive studies, and interviews, and quantitative content analysis of media coverage. Among the studies based on an historical approach with multiple qualitative data, Carpenter's (2001) work relied on comparative history and narrative panel analysis and compared three cases with narratives to analyze variations across three agencies (USDA, USPS, and the Department of the Interior) over time. Roberts (2006) also analyzed

FEMA's case with an historical approach to demonstrate the way the agency's reputation has influenced its autonomy. To analyze the temporal process that led to the FDA's jurisdictional claims, Maor (2010) used a historical-institutional analysis of four novel technologies (human gene therapy, lab-developed complex diagnostic tests, human tissue transplants, and human cloning experiments) based on an official journal of the U.S. government, policy announcements, and interviews. Moore (2015) conducted a case study of the VA based on an archival analysis of internal reports and private notes collected from the presidents' libraries, records of the American Association of Medical Colleges, and the VA.

Some studies have relied on multiple qualitative data, such as interview or archival studies. Busuioc's (2016) work relied on documents and 21 semi-structured interviews with expert EU and national-level respondents from Frontex and Europol. Maor (2007) analyzed the reputation protection mechanism of three drug reimbursement agencies in New Zealand, Australia, and British Columbia based on official documents, media coverage, and interviews.

Carpenter (2002) tested hypotheses using maximum-likelihood duration analysis to identify the factors that were associated with the FDA's drug review time between 1977 and 2000. Krause and Douglas (2005) compared three different types of agencies (Council of Economic Advisers, Office of Management of Budget, Congressional Budget Office, Federal Reserve) based on fiscal data to test whether reputation or political pressures drive agencies' performance. Luoma-aho (2008) conducted a survey of stakeholders in 12 public organizations in Finland to classify their reputation types based on semantic differences, including Authority, Trust, Service, Esteem, and Efficiency.

Further, some studies have relied on a systematic content analysis of documents. In his case study of Hurricane Katrina, Moynihan (2012) used systematic analysis of qualitative data in public reports from multiple government sources, such as the White House, U.S. House committee, U.S. Senate committee of Homeland Security and Government Affairs. His purpose was to explore the strategies that political actors used after Hurricane Katrina occurred. Christensen and Lodge (2016) focused on the way organizations seek to manage their reputations by accounting for their activities on their websites, and used content analysis of public agencies' websites related to intelligence, food safety, and flood defense.

Lastly, studies of bureaucratic reputation have used content analysis of media coverage. Maor (2011) employed an institutional analysis of procedures in the way public warning issues follow safety-based drug withdrawals for nine countries during 1975-2004. The author also analyzed media coverage to capture public warnings during the ten days following the withdrawal announcement. Maor et al. (2015) used media coverage as well to analyze the strategies that Israel's central bank used in response to its reputational threats. To test the effect of reputational threats on the FDA's enforcement activities, Maor and Sulitzeanu-Kenan (2013) used media coverage to measure the media's valence (positive, neutral, negative) and duration analysis of the duration of warning letters that the FDA issued. Similarly, Gilad, Maor, and Bloom (2013) used content analysis of media coverage to test the way the Israeli banking supervision department's response strategies varied according to their perceived reputational threats (media salience, media valence). Maor and Sulitzeanu-Kenan (2016) used a quantitative content analysis of media coverage in a

national newspaper in Australia to test the way reputational threats (valence) influenced government agencies' outputs.

3. Roadmap of the Dissertation

3.1 Gaps in the bureaucratic reputation literature

One of the main limitations to the theoretical approach in the bureaucratic reputation literature is that it focuses largely on the accountability-based approach to study reputations. This is no surprise, as bureaucratic reputation studies are rooted in political science, where researchers have major interests in the power and control between elected officials and bureaucrats. Nevertheless, as described in the previous section, reputation research also has great potential to enrich the public administration literature by applying network forms of governance, NPM, resource management, and communication strategies.

In addition, while the bureaucratic reputation literature acknowledges that multiple audiences formulate reputations according to different expectations, values, and motivations that influence the process, previous studies have focused only to a limited degree on citizens, despite their importance in a democracy. Among various audiences, previous studies have focused chiefly on politicians, the media, and other public organizations, in which the media were often equated with, or included in, media coverage. In particular, beliefs or opinions that citizens hold are associated with the legitimacy element of reputation, which shapes the power and autonomy that public agencies can obtain (Carpenter, 2010). Moreover, in this democratic society, citizens hold the position as the ultimate principal in public administration within the "chain of delegation," and as public agencies' accountability to citizens has become more important in today's society

because of its increasing complexity, citizens' evaluations also are assigned more weight. This is consistent with the shifting paradigm of public organizations that respond directly to citizens' assessments of service provision and quality (Byrkjeflot, 2015). Thus, this study focused on the citizens' side of bureaucratic reputation and the way it emerges through their cognitive processes.

Methodologically, one of the key limitations found in this field is that previous researchers have focused chiefly on organizational behavior at the macro-level. Historically, research has considered public organizations as the unit of analysis and studied their behaviors in response to their reputational concerns, such as response time and output changes. While studying bureaucratic reputation at the macro-level has benefits, as reputations imply audiences' collective views, and previous studies have provided many important findings that have contributed to the public administration and political science literature, this approach has certain limitations.

Firstly, the macro-level approach reveals the full mechanisms of bureaucratic reputation only in part. There are two major contrasting doctrines in the study of social phenomena: methodological individualism and methodological holism (Weber, 1922; Watkins, 1952). The goal of both approaches is to understand a social phenomenon; while the former emphasizes individual actions, the latter focuses on collective aspects. Both approaches have been debated historically, but this study adopted the methodological individualism view that the "macro" approach to social phenomena is incomplete without a "micro" foundation (Hayek, 1942; Alexander, 1987). Specifically, methodological individualism assumes that all social interactions are followed by interactions among

individuals (Arrow, 1994), indicating that psychological state, such as motivation or intentions, are important subjects.

Bureaucratic reputation is also a social phenomenon constructed by interactions between agencies and audiences, as Carpenter's (2010) definition states that it is "...a network of multiple audiences of organization." Further, Maor (2016) argued that agency-audience relationships are critical to an understanding of reputation formation's mechanism, for one of the assumptions bureaucratic reputation implies is that audiences are those who observe public agencies. He stated that audiences' perceptions are diverse, as different motivations, emotions, prior knowledge, and heuristics are involved in their views. This indicates that an approach that assumes that an organization is composed of persons, but studies their behavior collectively, can veil those important individual mechanisms. In fact, the purpose of this study, focusing on the "micro" foundation of bureaucratic reputation, is not to argue that individual-focused research is superior to collectivistic research, but that this approach is in need to achieve a complete understanding of bureaucratic reputation theory.

The second limitation lies in previous studies' methodologies, which have used largely qualitative approaches, including archival and historical research, interviews, and content analysis of media coverage. These methods have contributed to the literature by explaining reputation with respect to collective views at aggregate levels. Yet, as discussed above, those collectivistic methods provide more abstract explanations that lack discussion of individuals' psychological aspects, and what are needed now are methodologies that probe the determinants of bureaucratic reputation at an individual level. Thus, this study relied on a survey, survey experiment, and scale development, all of which have been

discussed as behavioral approaches to public administration (Grimmelikhuijsen, Jilke, Olsen, & Tummers, 2016). By focusing on micro-levels of bureaucratic reputation, this approach can enrich the reputation literature, as well as bridge the discussion of bureaucratic reputation to address the ‘Big Questions’ of public administrations (Jilke, 2016).

Some people might be concerned about micro-level approaches that use a survey or survey experiment, as they reveal only cross-sectional attributes of bureaucratic reputation, while reputation implies judgments within the interaction of organizations and audiences within the network that evolve historically (Chun, 2005). However, the goal of this study was not to prioritize individuals or provide “sufficient” answers to the questions related to bureaucratic reputations, but to provide one side of the “proof-of-concept” by examining behavioral implications (Jilke, 2016, p. 4). A snapshot of bureaucratic reputation within accumulated reputations enables us to study in-depth phenomena at a given moment, to which the current literature has given little attention, and which provides certain empirical findings that could link future studies to important research questions related to bureaucratic reputation.

3.2 Research purposes

The purpose of this study was to enhance our understanding of citizens’ reputation judgments of federal agencies in the U.S. Using an accountability-based approach, bureaucratic reputation is considered an external accountability mechanism in governance and focuses particularly on citizen-state interactions. While it focuses on citizens, this dissertation’s approach is consistent with NPM, which emphasizes the

performance-oriented regime, where governmental outcomes are considered critical aspects of accountability. This effort shed light on citizens' role as an external controller of administrative agencies in addition to elected officials, who pass judgments on, and pose questions about, an actor's conduct—in this case, governments (Bovens, 2007; Pollit, 2011; Osborne & Gabler, 1993; Van de Walla, 2016; Van Ryzin, 2004; 2013).

Specifically, while perception-based citizens' evaluations, such as their satisfaction or trust in government, have been considered a 'soft' measure of performance (Brudney & England, 1982), their ability to capture people's evaluations of organizations fully has limitations. Citizen satisfaction is very service-specific, while whether trust in government is associated with government-related factors is elusive (Bouckaert & Van de Walla, 2003). Further, Ostrom (1997) argued that a reputation for trustworthiness is a precursor of trust, which implies that trust is the accumulation of reputation judgments from repeated interactions among actors. This study suggests that organizational reputation is people's judgments of public organizations, which captures more than outcomes, but also incorporates their morality and procedures comprehensively (U.S. Congress, Committee on Governmental Affairs, section 2).

In addition, there has been a recent debate over the definition of the term "citizen" in public administration (Roberts, 2020). In academia, citizens have been defined in various ways, depending on their legal status (Cohen and Ghosh, 2019), or whether they are residents (Robinson et al., 2019) or inhabitant (Stevenson, 2010). In this study, the term citizen is defined as an adult who resides in the U.S and thus is part of the society served by administrative agencies, which has been the common approach of research in

public administration. Also, the definition could be justified that the purpose of this study is to view the general attitudes of the U.S. population towards federal agencies.

Grounded in this framework, this study attempted firstly to overcome the problem of the dominant macro-level approach in the bureaucratic reputation literature by providing relevant survey tools that can capture citizens' reputation judgments, and which also can be used to measure that of stakeholders. This study was undertaken in recognition that the current literature has no way to measure individual-level reputation despite its need and importance. A bureaucratic reputation scale could be a foundation on which to build future studies to investigate more of the 'big questions' related to the topic. For example, using such a scale can determine both the ways that account holders and holdees interact using reputations as signals and the way organizations can formulate and manage their reputation, which is a crucial resource.

3.3 Research questions

Q1. How can we measure bureaucratic reputation at the individual level?

Study 1's purpose was to develop a standardized measure that can be used in individual-level studies of bureaucratic reputation. Surveys or survey experiments that focus on bureaucratic reputation as a key variable could benefit from a standardized measurement of bureaucratic reputation that is applicable to various types of agencies and levels of government. As a result, Study1 followed DeVillis' (2003) standard scale development processes: developing a large item pool from previous literature; refining the items' content validity with the help of experts; using exploratory and confirmatory factor

analyses to identify a short subset of items to form the scale, and examining the reliability of the scale as well as its empirical validity.

Q2. What factors influence citizens' reputation judgments?

The purpose of Study 2 was to examine the factors that are associated with the formation of citizens' reputation judgments of U.S. federal agencies. Study 2 had an exploratory purpose and used secondary data from the Pew Research Center that focuses on the way citizens' characteristics and contexts are associated with their reputation judgments of 12 U.S. federal agencies. Using the theoretical framework of policy preference, this study tested whether reputation judgments are related to individuals' demographics (race, age, gender), socioeconomic status (income and education), geographic characteristics (Northeast, West, South, Midwest/Urban, suburban, rural), and/or ideologies (political ideology, trust in government).

Q3. How can public organizations manage citizens' reputation judgments?

Study 3's purpose was to probe ways that public organizations can manage their reputation using various communication strategies. While many studies have focused on how reputational concerns shape agencies' selection of communication strategies, few have discussed the way citizens respond to those strategies. Thus, Study 3 conducted a survey experiment to test citizens' responses to different blame-avoidance strategies government agencies typically employ. In particular, the study focused on the effectiveness of admission, denial, and nonresponse strategies. The study also tested whether citizens' responses depend on their prior judgments of agencies' reputations with the theoretical

framework of a halo effect that focuses on individuals' cognitive processes. The study recruited approximately 1,000 respondents through Qualtrics.

Note on the publication status and authorship of studies in this dissertation

This dissertation is composed of three essays (or papers), two of which (Chapter 2 and Chapter 3) have already been published in peer-reviewed journals. Chapter 2 was published in *Governance* (Lee & Van Ryzin, 2019), and Chapter 3 was published in the *International Review of Administrative Science* (Lee & Van Ryzin, 2020). Although Dr. Gregg Van Ryzin (chair of this dissertation) was a second author on these published studies, the original ideas for both papers as well as the data collection, statistical analysis, and initial writing were completed by Danbee Lee as lead author. Dr. Van Ryzin provided guidance on the research designs of both studies, suggested edits to the initial draft papers, and helped in responding to the anonymous peer reviews. The PhD Committee at Rutgers SPAA was notified in advance about the co-authorship status of both these papers and approved their inclusion in this dissertation. In addition, this dissertation followed the *Guidelines on Using Previously Published Work in Theses and Dissertations*, endorsed by the Rutgers School of Graduate Studies in March 2018. This includes determination that the publication agreements of both Wiley (*Governance*) and SAGE (*IRAS*) allow for use of the work in a thesis or dissertation without obtaining written permission of the publisher. Thus, the pre-publication versions of both papers are included, with minor edits to accommodate the dissertation format, as Chapter 2 and Chapter 3. Chapter 4 is a separate experimental study that has not yet been published. Although Dr. Van Ryzin provided

guidance on the design of the experiments, Danbee Lee is the principal investigator and sole author of Chapter 4.

CHAPTER 2:

MEASURING BUREAUCRATIC REPUTATION

From: Lee, D., & Van Ryzin, G. G. (2019). Measuring bureaucratic reputation: Scale development and validation. *Governance*, 32(1), 177-192.

1. Summary

Despite growing interest in bureaucratic reputation as a theoretical construct, the field lacks a standardized measure that can be used in surveys to capture individual-level variation in the reputation judgments of citizens and other audiences. The aim of this study, therefore, is to develop a standardized, individual-level measure of bureaucratic reputation based on the conceptual definition provided by Carpenter (2010). Employing feedback from experts and data from a survey of over 300 U.S. citizens, this study develops and tests a unidimensional scale of bureaucratic reputation, representing the content domains of performance, morality, procedural fairness, technical competence, and general reputation. Results suggest that our proposed Bureaucratic Reputation Scale (BRS) has good internal reliability and that it is positively associated with support for autonomy, budget, and power, which provides evidence of criterion validity. Potential uses of the scale to study bureaucratic reputation are discussed.

2. Introduction

Interest in the theory of bureaucratic reputation has increased in the last decade with a body of consistent and robust findings showing strategic behaviors of public agencies to enhance their reputations while protecting themselves from reputational

threats (Carpenter 2010; Moynihan 2012; Gilad, Maor, and Bloom 2015). Bureaucratic reputation has been defined as a set of beliefs on the part of various audiences about an agency's unique capacities, roles, and obligations (Carpenter 2010). Reputation provides a valuable resource for public organizations because of its connections to an organization's power, autonomy, and legitimacy (Carpenter 2001, 2010; Maor 2007). Despite growing interest in the concept and its theoretical importance, little attention has been given to measuring bureaucratic reputation systematically. Previous research on bureaucratic reputation has relied on mostly qualitative approaches, including case studies, archival data, interviews, and content analysis of media coverage (Carpenter 2001; Maor 2016). These methodological approaches are important and insightful, especially for studying reputation at macro levels. However, research on bureaucratic reputation could also benefit from standardized measurement at microlevels, focusing on individual citizens and other audiences. Thus, this study reports on the development of a short 5-item scale and a longer 10-item scale that measures bureaucratic reputation, and that can be used for various types of agencies and audiences, including citizens.

3. Research Background

3.1 The concept of bureaucratic reputation

Reputation is a broad concept that appears in organizational theory in economics, sociology, and business, with varying definitions (Chun 2005; Fombrun, Gardberg, and Sever 2000; Schwaiger 2004). For example, in the business world, Fombrun et al. (2000, 242) have defined reputation as “a collective construct that describes the aggregate perceptions of multiple stakeholders about a company's performance.” We ground our

work, however, on Carpenter's (2010) definition, which focuses on the reputation of public organizations; he defines bureaucratic reputation as "a set of symbolic beliefs about the unique or separable capacities, roles, and obligations of an organization, where these beliefs are embedded in audience networks" (Carpenter 2010, 45). This definition of reputation emphasizes, "the evaluation of the organization's unique character and activities by multiple audiences" (Maor 2016, 82). While highlighting the collective judgment of multiple audiences, bureaucratic reputation does not focus solely on performance, as in much of the business and economics literature, but also other dimensions that are especially important in a democratic society (such as morality and procedural justice), as discussed shortly. Thus, in Carpenter's formulation, bureaucratic reputation possesses both multiplicity and subjectivity, meaning that it is related to multiple attributes of an organization and to various audiences.

Specifically, bureaucratic reputation, as conceptualized by Carpenter (2010), includes the content domains of performance, morality, procedure, and technical competence, which together form a public organization's overall reputation. *Performance reputation* refers to how audiences evaluate the organization's "quality of decision making," "capacity for effectively achieving its ends," or "announced objectives" (Carpenter 2010, 46). Carpenter (2001) suggests that this content domain is the most fundamental aspect of bureaucratic reputation. Performance reputation can also refer to the audience's perception of an agency's "vigor and aggressiveness" when pursuing goals. Performance reputation operates by "inviting compliance, inducing decision that renders the agency's work easier or less controversial or deterring challenges to the

organization's power" (Carpenter 2010, 46). Busuioc and Lodge (2016) point out that performance reputation is enhanced by achieving popular policy outputs and outcomes.

Moral reputation deals with the values and ethics of an organization as recognized by key audiences. From a legal or constitutional perspective, scholars have argued that public administration has special moral obligations (Hart 1984; Rohr 1988). Morality and ethics are essential normative components of how public officials should conduct themselves, including the ethical behavior of honesty, conformity to law, and treating individuals and groups fairly based on an understanding of regime values (Willbern 1984; Hart 1984). Also, Rohr (1988) argues that a public administrator's primary moral obligation is to protect the founding values of democracy. Thus, moral reputation reflects an organization's ethical behavior, including its honesty, integrity, and conformity to law and social norms.

Procedural reputation refers to the "justness of the processes" of the organization and its norms of deliberation, procedure, or decision-making (Carpenter 2010, 47). In a complex society where diverse interests conflict, audiences often judge an agency based on the fairness of the processes it employs to reach decisions, even if the outcomes do not satisfy everyone (Van Ryzin 2011; Willbern 1984). Following Busuioc and Lodge (2016), procedural reputation is about following the "right" rules in a given situation. Thus, procedural reputation concerns whether the processes the public agencies conduct are fair and reliable. While procedural reputation and moral reputation clearly overlap, Carpenter (2010, 47) suggests they are different to the extent that an organization "may have defensible aims and ethically appropriate strategies for meeting them, but may not have followed commonly recognized norms of deliberation, procedure, or decision

making.” Thus, an agency may have moral ends, but procedurally deficient means; alternatively, an agency can have just means (fair procedures), but perhaps morally questionable ends.

Technical reputation concerns an organization’s scientific authority, methodological sophistication, and analytical capacity (Carpenter 2010, 47). As Carpenter details in his monumental history of the US Food and Drug Administration (FDA), the FDA’s reputation rested to a large extent on the agency’s scientific and technical authority in matters of drug testing and in establishing evidence-based standards for safety and efficacy. Specialization and technical knowledge also core professional values of bureaucrats (Friedrich 1940; Perry 1997). Busuioc and Lodge (2016) also point out that technical reputation relates closely to an agency’s subject expertise, such as the FDA’s expertise in drug testing. In this way, technical reputation emphasizes a particular body of knowledge, skills, and experience when public agencies address technical issues over which they have jurisdiction.

In addition to the specific content domains of bureaucratic reputation discussed above, general attitudes and feeling that various audiences have towards an organization also play a role (Fombrun et al. 2000). The division of emotional and rational aspects of reputation is well-established in the business literature, with the emotional aspect often linked to an overall evaluation of reputation (Fombrun et al. 2000; Schwaiger 2004). For example, people often possess a general disposition toward an organization based on a “personal appreciation of how a company is and what it does” or affective judgment (Schwaiger 2004, 56). Thus, in our conceptualization, a general reputation domain is included that reflects people’s overall feelings or attitudes towards an agency.

3.2 The measurement of reputation

Previous efforts to quantify reputation have focused mostly on business organizations and corporate reputation. Indeed, business magazines and media firms regularly conduct various surveys to measure corporate reputations, including *Fortune*, *Manager Magazine*, *Management Today*, *Asian Business*, *Far Eastern Economic Review*, *Financial Times*, and *Industry Week* (Fombrun et al. 2000). Additionally, business researchers have developed various standardized scales, including the reputation quotient (RQ) model, the customer-based corporate reputation model, the resource-based view model, and the personification metaphor model (Fombrun et al. 2000; Davies, Chun, Silva, and Roper 2001; Boyd, Bergh, and Kitchen Jr. 2010; Walsh and Betty 2007). Of course, with the rise of online commerce, the reputations of businesses are also measured by various consumer ratings from websites such as Google, Amazon, Yelp, and others. Interest in measuring organizational reputation has also been growing in nonprofit management research. For example, Bennett and Gabriel (2003) developed a measure of the reputation of UK charities by factoring semantic differences describing the characteristics of organizations. Also, Sarstedt and Schloderer (2010) developed a measurement of the reputation of nonprofit organizations using scales and dimensions originally established by Schwaiger (2004) for corporate reputation. There are also the published ratings of nonprofits provided by watchdog organizations such as Charity Navigator and GuideStar.

Compared with business and nonprofit management, public management as a field has devoted less attention to developing a standardized measurement of reputation. Carpenter's (2010) seminal work mostly employed in-depth historical research and case

studies to investigate the reputation of federal agencies. Others used content analysis of media coverage to track bureaucratic reputation (Maor 2011; Gilad et al. 2013; Maor and Sulitzeanu-Kenan 2015). Picci (2011) suggested measuring government reputation by calculating a reputation index consisting of the aggregated assessment of citizens over time as a way to provide signals and incentives for better government performance, although such indexes have yet to be widely implemented. To date, only a few studies in the field have attempted to gauge bureaucratic reputation judgments at the level of individual citizens or audience members. Luoma-ho (2008) developed a measure based on semantic differentials by surveying the agency's multiple audiences. This attempt has enhanced the understanding of reputation at micro-levels; however, the approach has some limitations. The method of semantic differentiation captures the emotional aspect of reputation, while this approach misses important attributes of bureaucratic reputation that involve both subjectivity and multiplicity (Carpenter 2010).

To address the need for a more comprehensive and standardized measure, the goal of this study was to develop and test a multi-item scale of bureaucratic reputation, grounded in Carpenter's (2010) theory, that can be used with various types of public agencies and their audiences. The scale seeks to represent the key content domains of reputation identified by Carpenter (2010)—performance, morality, procedure, and technical ability—plus the general, a more affective aspect of reputation. This study attempted to develop a scale that would be versatile and practical so that it can be used in various types of surveys or other studies to investigate bureaucratic reputation as perceived by citizens or other audiences. Of course, all survey measures have limitations related to the problem of self-reporting and the inherent common source bias that confound much of the analysis done

with survey data. Still, survey research remains a fundamental tool of social research and benefits from the incorporation of standardized, validated measures of key theoretical concepts.

4. Scale Development Processes

4.1 Item generation

To begin with, an initial pool of agree-disagree items was generated from a review of the reputation literature. The review process included several items from previous measurements of the reputation of for-profit and non-profit organizations, including the studies discussed above by Fombrun et al. (2000), Schwaiger (2004), and Sarstedt and Schloderer (2010). However, many of the items from these scales could not be used or even adapted because they focused on share prices, profits, investments, and other features not relevant to the public sector (Maor 2016). Thus, many of our own agree-disagree items were devised based on statements suggested in Carpenter's (2001, 2010) work, specifically with respect to his discussions of performance, morality, procedure, and technical competence as core content domains of bureaucratic reputation. Additionally, items were developed from the writing of other scholars who have defined and measured bureaucratic reputation (Busuioc and Lodge 2016; Luoma-ho 2008; Moar 2016; Picci 2011). Items and statements were consulted with colleagues and those ambiguous, double-barreled, or redundant were removed. By the end, an initial pool of 46 items were produced representing five content domains of bureaucratic reputation: performance, morality, procedure, and technical competence, and general reputation.

4.2 Expert review

The next step was to have our initial item pool reviewed and refined by independent experts. The experts were scholars of public administration and political science who had published articles on bureaucratic reputation in peer-reviewed journals in many different countries around the world, which would support the broad applicability of our scale. In the end, feedback on the item pool were identified and obtained from 22 experts, who were sent an online survey that presented our 46 items and asked for feedback and suggestions on each item. Specifically, in the body of the survey, the experts were first given conceptual descriptions of the five domains of bureaucratic reputation and asked to evaluate each of the 46 items as an “excellent,” “good,” “fair,” or “poor” representation of its domain. Experts were also asked for any additional open-ended comments as well as specific suggestions for revising any of the items. Based on the results of the 22 experts’ ratings and remarks, items were removed that the experts consistently rated less than “good,” reducing the 46 items in our pool down to 30 items, with each dimension having six items. In addition, the open-ended comments from the experts were used to revise the wording of some of the remaining items. Additional attempts were made to reframe some of the items so that each dimension had some negatively worded items as well as positively worded items.

4.3 Citizen survey

An online survey of citizens was conducted to gather data to evaluate the reliability and validity of the scale. As note earlier, the theory of bureaucratic reputation holds that reputation is composed of the views of multiple audiences, including elected

officials, media, interest groups, and citizens (Carpenter and Krause, 2012). Although our scale is tested only on a sample of citizens, this limitation can be justified to the extent that citizens are often seen as the ultimate principal in democratic government and by the importance of public opinion in the political process. Respondents came from an online research panel composed of adult volunteers recruited from across the U.S. and other countries; the panel (CivicPanel.org) is used for academic research and is housed at Rutgers University's School of Public Affairs and Administration (see <https://spaa.newark.rutgers.edu/civicpanel>). For purposes of this study, the survey invitations were sent to U.S. participants only. The data collection process started on April 14 and ended on May 8, 2017. A total of 7,819 panel members received an email invitation, of which 1,183 opened the email invitation. A total of 419 people responded after two contact attempts (representing a 15% contact rate and 35% cooperation rate). After removing observations having more than one missing value on the core items, the total respondents came to 348. Appendix 1 describes the respondents' characteristics, including gender, race, age, education, income, location, and political ideology. About two-thirds were female, 84% were white, and 45% of the respondents were 55 or older.

The survey began by asking respondents to "indicate the extent to which you agree or disagree with these statements about the [name of agency]." Here we randomly inserted the name of one of the following three federal agencies: The Food and Drug Administration (FDA), National Aeronautics and Space Administration (NASA), and the Internal Revenue Service (IRS). These agencies were selected because they represented agencies with high, medium, and low favorability ratings according to a Pew Research Center (2013) survey, one of the few publicly available surveys that gauge opinions about

specific federal agencies (for more information, see <http://pewrsr.ch/19XpU9k>). The statements in our survey of citizens were the 30 items that was developed from previous review of the literature and from the feedback of the 22 experts (as described above). 30 items were scrambled and then divided the into three blocks, without ordering or framing the items in any way to indicate their content domains. Thus, the order of items within blocks was randomized, and the order of the blocks was randomized, for each respondent. Respondents rated their level of agreement with the statements on a 5-point scale (from 1=*Strongly disagree*, to 5=*Strongly agree*). In this way, attempts were made to make sure that the content (wording) of the items and not the presentation of the items would determine their inter-correlations.

4.4 Factor analysis and reliability test

Analytical steps began with exploratory factor analysis of the 30 items using principal component factoring with orthogonal rotation, an approach that is well suited to scale development (Devellis 2003; Acock 2012). One random split-half sample (n=163) were used to run exploratory factor analysis, and reserved the remaining split-half sample for running confirmatory factor analysis (described later). Table 2.1 presents the results of the exploratory factor analysis, showing a three-factor solution and the item-factor loadings. The first factor is clearly dominant, with the most high-loading items and accounting for fully 34 percent of the total variance in all items. Originally included was both positively and negatively worded items in the 30-item pool in an attempt to help avoid acquiescence or agreement bias (Devellis 2003) but, unfortunately, the negatively worded items all loaded on a separate factor (factor 2). Thus, although these results

suggest perhaps that negative reputation is a different dimension from positive reputation, as trust and distrust are sometimes viewed as separate dimensions (McKnight, Choudbury, and Kacmar, 2002), items from factor 1 were chosen to represent the core content of our scale.

As a robustness check, the same exploratory factor analysis was ran using multiple split-half samples (Devellis 2003). That is, our sample was randomly divided in half five times and then re-ran the exploratory factor analysis separately for each split-half sample (see Appendix 2). This approach tests the robustness of the factor solution and highlights items that consistently loaded on the dominant factor; it was the first factor in all 5 split-half samples, which suggests they are fairly robust indicators of the latent construct.

Table 2.1 Exploratory Factor Analysis

Variable		Mean	SD	Factor 1	Factor 2	Factor 3
item1	I have a favorable opinion about this agency	3.48	1.17	0.68		
item2	I believe this agency is doing a good job	3.52	1.12	0.71		
item3	I don't have much respect for this agency (R)	3.47	1.33		0.87	
item4	Overall, this agency has a good reputation	3.48	1.18	0.63		
item5	This agency is a waste of taxpayer's money (R)	3.47	1.34		0.82	
item6	I have a negative impression about this agency (R)	3.37	1.37		0.84	
item7	This agency is a well-run organization	3.44	1.08	0.69		
item8	This agency is effective at its job	3.60	1.05	0.62		
item9	This agency does a poor job (R)	3.44	1.32		0.85	
item10	This agency is a high performing agency	3.56	1.05	0.68		
item11	This agency often fails to get things done (R)	3.06	1.19		0.73	
item12	This agency has the capacity to get things done	3.84	1.05	0.58		0.56
item13	This agency can be trusted to do what is right	3.51	1.14	0.60		0.61
item14	This agency maintains high ethical standards	3.53	1.10	0.66		0.51
item15	This agency seems to be corrupt (R)	3.47	1.25		0.72	
item16	This agency protects democratic values	3.37	0.96	0.61		
item17	This agency is sometimes dishonest (R)	3.21	1.24		0.71	
item18	I believe what this agency says	3.41	1.10	0.74		
item19	This agency respects due process	3.50	1.01	0.81		
item20	This agency is highly transparent	3.10	1.08	0.75		
item21	Although I sometimes disagree with its decisions, this agency always follows the rules	3.29	1.01	0.78		
item22	This agency treats people fairly	3.40	1.07	0.85		
item23	This agency protects the rights of citizens	3.37	1.03	0.80		
item24	This agency is politically neutral	3.19	1.03	0.67		
item25	This agency has the technical expertise to do its job well	3.77	1.07			0.76
item26	This agency is technically competent	3.83	1.00			0.72
item27	This agency has highly skilled employees	3.85	1.07			0.66
item28	This agency bases its decisions on evidence	3.64	1.04	0.71		
item29	This agency has the skill to deal with complex situations	3.73	1.05	0.63		
item30	This agency lacks technical knowledge (R)	3.64	1.25		0.77	
Percentage of variance (%)				34	23	17

Note: N=163, principal components factoring (pcf) in Stata with varimax rotation. Bolded factor loadings indicate items selected for the scale. Factor loadings < 0.5 are not shown. R = reversed item.

Using the results of our exploratory factor analysis, including the split-half robustness checks, items were selected that had high factor loadings and that were relatively stable across split-half samples. Also, items were chosen that were conceptually distinct from each other and that represented a balance of items across the five content domains of bureaucratic reputation. Thus, a combination of statistical criteria (from the factor analyses) and substantive judgment (based on the theory of bureaucratic reputation) were used in the selection of items for the final scale. The selected items are shown in bold in Table 2.1. A reliability analyses were ran for both a 5-item short scale and a 10-item longer scale using Cronbach's alpha, which is defined as "the proportion of variance attributable to the true score of the latent variable" and is the most widely used and conservative reliability estimate (Devellis 2003, 31; see also Carmines and Zeller 1979). For the 5-item scale, $\alpha = .92$; and for the 10-item scale, $\alpha = .96$. Thus, both short and long scales demonstrate a high degree of internal consistency.

To provide additional tests of reliability, the reserved split-half sample (n=174) was used, and it generated model fit indices from confirmatory factor analysis (CFA). Using CFA after EFA is a common approach to confirm the statistical reliability and validity of a scale (Worthington and Whittaker 2006). Figure 2.1 shows the result of our 5-item scale. The fit statistics are chi-square = 8.33 ($p > 0.01$), the root mean square error of approximation RMSEA=0.06, a comparative fit index CFI=0.99, and a standardized root mean squared residual SRMR=0.02, which all indicate an acceptable fit of the data to the unidimensional model. Also, all of the items significantly load on the latent variable, with standardized coefficients ranging from .81 to .88, which indicate fairly strong relationships

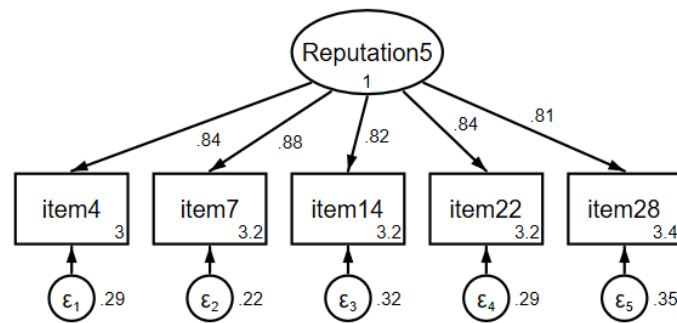


Figure 2.1: Confirmatory Factor Analysis for 5-item Scale (Standardized)

Figure 2.2 shows the CFA results of our 10-item scale. The fit statistics are chi-square = 134.73 ($p < 0.01$), RMSEA=0.13, CFI=0.94, and SRMR=0.04. The CFI and

SRMR indicate a very good fit of the data to the unidimensional model, however, the p-value of the chi-square is significant, and the RMSEA is above the standard threshold. The significant chi-square result can be explained in part by the increasing misfit that inevitably occurs when more items and thus parameters are added to a measurement model (Schermele-Engel, Moosbrugger, and Müller 2003). The RMSEA value of .13 is just above the standard .10 cut-off for mediocre fit; but such thresholds are somewhat arbitrary, and other studies have used less stringent standards (Chen, Curran, Bollen, Kirby, and Paxton 2008; Grimmelihijsen and Knies 2017). All 10 items had significant factor loadings, with standardized coefficients ranging from 0.7 to 0.89. Still, these results suggest that the overall fit of the 10-item scale is not as good as the 5-item scale.

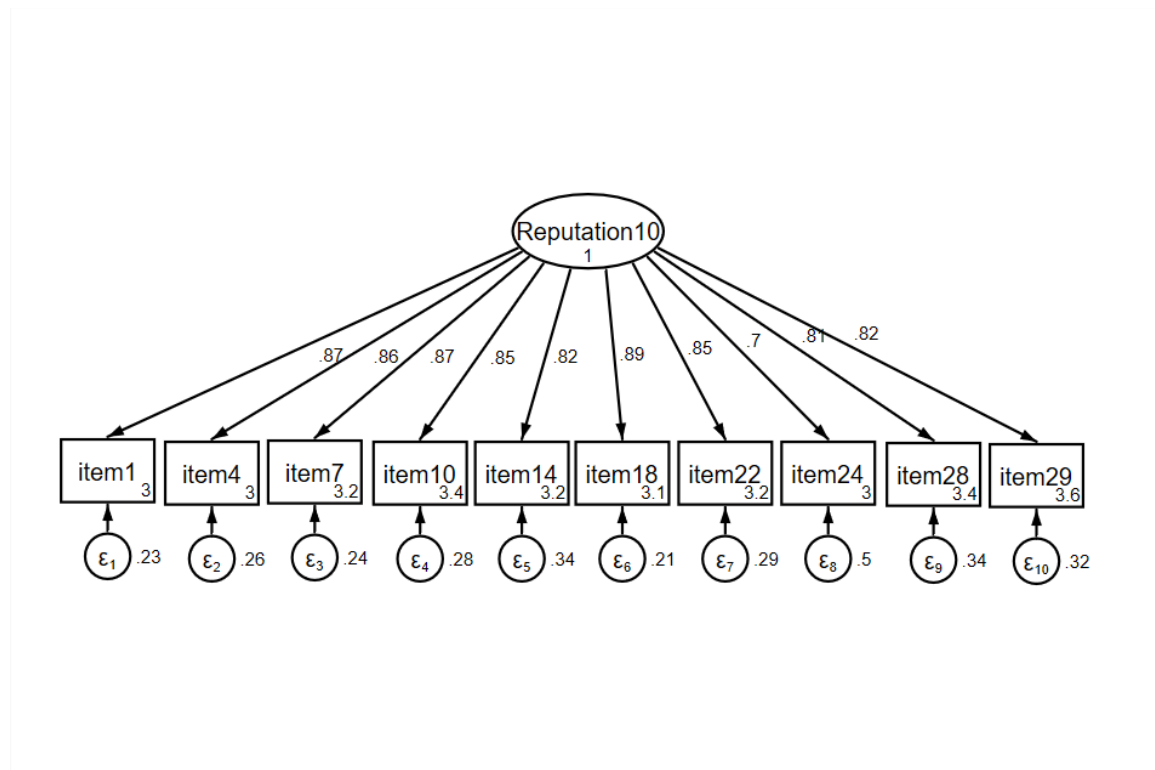


Figure 2.2: Confirmatory Factor Analysis for 10-item Scale (Standardized)

4.5. Construct validity

Construct validity refers to how the established measure relates to other measures as theoretically expected (Carmines and Zeller 1979). In order to test construct validity, this study focused on convergent validity and discriminant validity; the former emphasizes the relationship between the measure and theoretically related constructs, while the latter focuses on the absence of the relationship between the measure and theoretically more distant or unrelated constructs (Campbell and Fiske 1959; Devilles 2003). This study followed three distinct steps of construct validity suggested by Carmines and Zeller (1979): specifying the theoretical relationship between the measure and concepts; examining these relationships empirically; and interpreting the evidence for construct validity. Thus, this study examined the association of our 5-item and 10-item bureaucratic reputation scales with a set of criterion variables that, according to bureaucratic reputation theory, were expected to be strongly associated: autonomy, budget, power, job performance, and favorability. Table 2.2 shows the bivariate correlations between our 5-item and 10-item bureaucratic reputation scales and these criterion variables, which are discussed in more detail below.

Table 2.2: Correlation analysis

	Mean	SD	Min- Max	1	2	3	4	5	6	7	8	9	10	11
1. reputation (5-item)	3.40	.94	1-5	1										
2. reputation (10-item)	3.40	.91	1-5	.98**	1									
3. giving more autonomy	3.02	1.12	1-5	.58**	.58**	1								
4. increasing budget	3.13	1.22	1-5	.60**	.61**	.73**	1							
5. favorability	2.90	.88	1-4	.84**	.84**	.58**	.62**	1						
6. job performance	2.78	.88	1-4	.86**	.86**	.56**	.59**	.85**	1					
7. power possession (underpowered vs. overpowered)	.66	.47	0-1	.46**	.46**	.36**	.41**	.46**	.44**	1				
8. power use (responsible vs. abusing)	.71	.45	0-1	.65**	.66**	.53**	.51**	.68**	.67**	.53**	1			
9. political ideology	3.04	1.28	1-5	-.08	-.10*	-.09	-.13**	-.07	-.08	-.23**	-.11**	1		
10. age	15.16	1.39	12-19	-.13	-.03	-.11*	-.06	-.02	-.03	.08	-.02	.04	1	
11. income	6.48	3.30	1-12	-.00	-.01	1.00*	.18*	.01	.03	-.08	-.01	-.02	-.03	1

SD=Standard Deviation; *p<0.10 **p<0.05 (two-tailed tests)

Bureaucratic *autonomy* is theorized to be dependent on reputation (Carpenter 2001; Robert 2006; Maor 2011). To measure autonomy, this study assessed respondents' agreement with the statement: "The [name of agency] should be given more autonomy" (1=Strongly disagree, to 5=Strongly agree). As Table 2.2 shows, both the 5-item and 10-item reputation scales are positively associated with support for increased autonomy. Reputation is also presumed to lead to support for *budget* increases, and indeed previous research has found a positive relationship between public opinion about agencies and support for more spending (Meier 1992). Thus, respondents' agreement were assessed with this statement: "The budget of the [name of agency] should be increased" (1=Strongly disagree, to 5=Strongly agree). As Table 2.2 shows, both the 5-item and 10-item reputation scales are positively related to support for budget increases. Granting additional *power* to an agency is another theorized benefit of bureaucratic reputation, as Carpenter's (2010) history of the FDA suggests. Two questions were included about the agency's power, which were originally used by Gallup. The first one is about the responsible *power use* of the agency: "Thinking about the powers the [name of agency] has been given to do its job, do you think [name of agency] generally uses powers responsibly (coded 1), or frequently abuses powers (coded 0). The other question is about the amount of *power possessed* by the agency: "The [name of agency] had been given the right amount of power, more power than it needs, too little power." Responses were dichotomized to this question to create a dummy variable (1=too little power/right amount, 0=more power than it needs). In accordance with theoretical expectations, results show that both the 5-item and 10-item bureaucratic reputation scales are positively associated with the perceived responsible use of power and support for granting an agency more power.

In addition to autonomy, budget, and power, favorability and job performance were also examined, which are based on standard survey questions used by the Pew Research Center and Gallup. Because reputation in part reflects overall attractiveness, likeability, and admiration of the organization (Fombrun 2012; Wæraas and Byrkjeflot 2012), as discussed earlier, this study would expect the 5-item and 10-item bureaucratic reputation scales to be positively related to favorability. As mentioned, Pew results show varied favorability scores, with NASA as the highest, the FDA in the middle, and the IRS among the lowest. The Pew question about *favorability* asks: “Is your overall opinion of the [name of agency] . . . (1=strongly favorable, to 4=strongly unfavorable). As Table 2.2 shows, both the 5-item and 10-item bureaucratic reputation scales are positively correlated with favorability. Performance is a key domain of reputation, as discussed earlier, so it makes sense that our proposed scales should be related to Gallup’s standard question about *job performance* of government agencies. Moreover, many empirical studies have found a relationship between performance and reputation, although these concepts clearly overlap, and the direction of causality remains ambiguous (Gilad et al. 2013; Krause and Douglas 2005; Maor and Sulitzeanu-Kenan 2015). The standard question used by Gallup is as follows: “How would you rate the job being done by [name of agency]?” (1=poor, to 4=excellent). As Table 2.2 shows, both the 5-item and 10-item bureaucratic reputation scales are positively correlated with job performance.

In addition to correlating with theoretically important criterion variables, a valid measure should *not* reflect extraneous or spurious influences from theoretically distant or unrelated variables. This is the notion of discriminant validity, which this study attempted to test with political ideology, age, and income. This study chose these three variables

because they are standard socioeconomic factors that can influence public opinion about government. With respect to political ideology, previous research has suggested that conservatives generally have a more negative view of government than liberals (Goodsell 2012). Indeed, a small negative correlation were found between political ideology (being more conservative) and our bureaucratic reputation scales, although the correlation is weak and only marginally significant for just the 10-item scale. Also, *age* was tested because cohort and life-stage effects often play a role in attitudes toward government (Christensen and Laegreid 2005). Our results show that age is not significantly correlated with either the 5-item or 10-item bureaucratic reputation scales. Lastly, *income* was analyzed because some literature has suggested that socioeconomic status and class differences shape attitudes towards government (Alford 2001). But as Table 2.2 shows, neither the 5-item scale nor the 10-item scale is significantly correlated with income. In sum, these results help confirm the discriminant validity of the reputation scales by showing their statistical independence from ideology, age, and income.

5. Agency Levels and Profiles

As explained earlier, our survey of citizens used three federal agencies as the referents of the test items, the FDA, NASA, and IRS. Respondents were allocated one of these agencies randomly and answered all 30 statements with the randomly allocated agency in mind. Thus, using our 5-item scale, we can compare the mean reputation ratings among these three federal agencies, as shown in Figure 2.3 NASA had the highest overall score (3.88), the FDA next highest (3.23), and the IRS the lowest (3.14). This rank parallels

the favorability ratings of these agencies from the 2013 Pew survey, providing further evidence of the validity of our proposed scale.

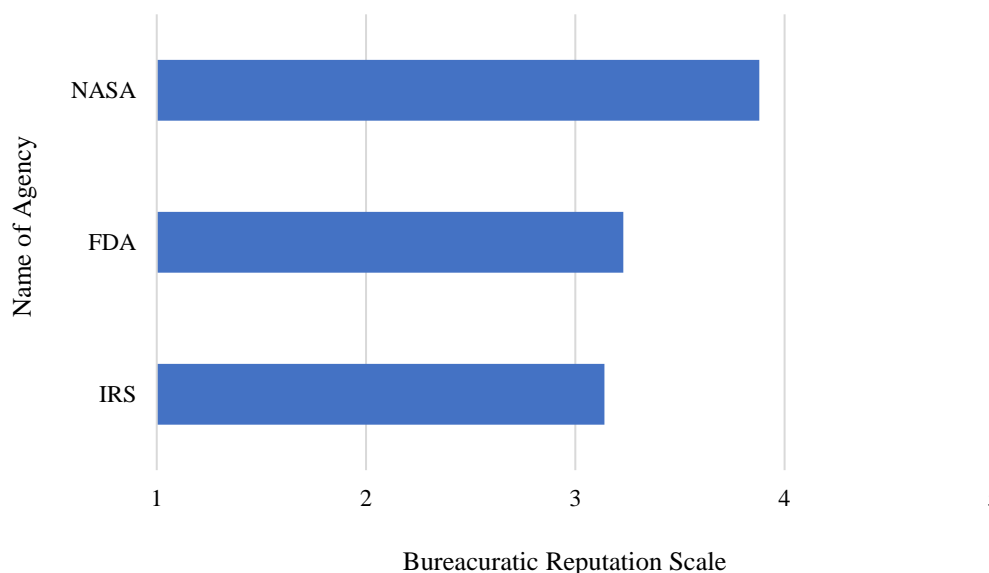


Figure 2.3: 5-items Reputation Scale Score

Figure 2.4 provides a visualization of the reputation profiles of the FDA, NASA and the IRS based on the 10-item reputation scale, which provides two items for each of the five content domains. Interestingly, the IRS has a lower general reputation (2.98) compared with the other four domains, meaning that people's general attitude towards the agency remains relatively negative. In contrast, NASA has a relatively higher general reputation rating (4.11), yet its procedural reputation score appears somewhat low (3.54). These results imply that each agency has a unique reputation profile that is not apparent when measuring overall reputation. For example, the IRS has a higher performance reputation than FDA, even though the FDA has a higher reputation score overall. This study presents these reputation profiles simply to illustrate the scale scores for each agency

in more detail, but such profiles could be potentially of substantive interest in their own right in future research.

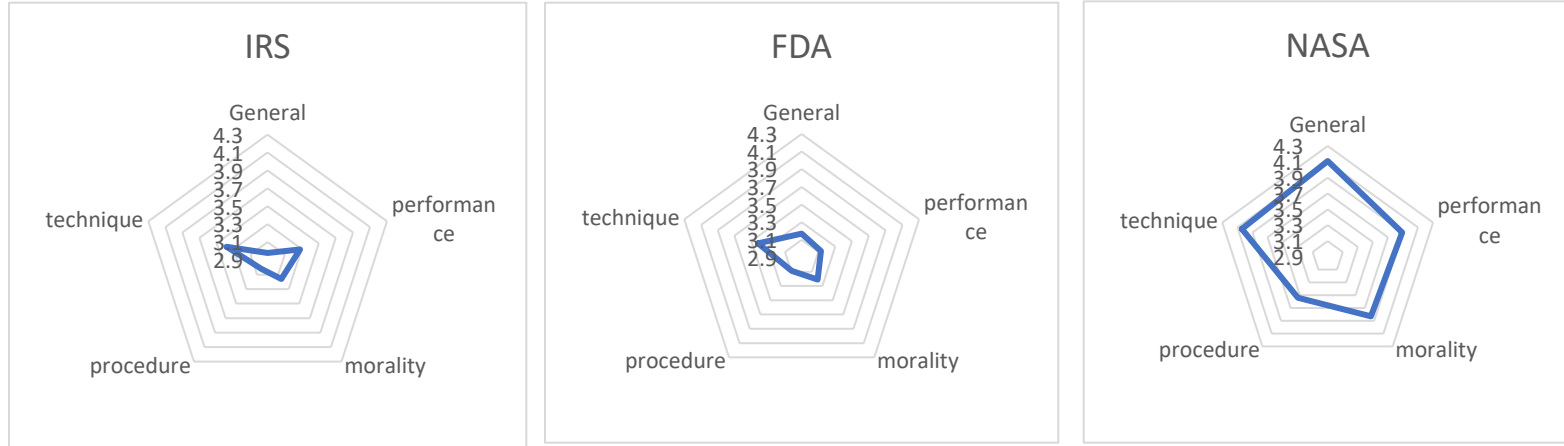


Figure 2.4: Dimensional Profile of Reputation Measure

6. Conclusions

To develop and test a standardized measurement of bureaucratic reputation, this study followed a scale development process that involved the theory-based generation of a large item pool, expert review and refinement of items, and the factor analysis of a nationwide citizen survey to empirically identify a short 5-item and a longer 10-item unidimensional scales of bureaucratic reputation. Both versions of the scale correlate strongly with support for agency autonomy, budget increases, and power, providing evidence of the theory-based construct validity of the scale. Both scales also correlated positively (as expected) with established polling questions that ask about the favorability and job performance of federal agencies. Additionally, this study found that both the short and long scales are largely statistically independent of political ideology, age, and income, providing evidence of discriminant validity. The final Bureaucratic Reputation Scale, both the short and long versions along with the instructions and the agree-disagree response format, are presented in Appendix 3 for use by other researchers.

Although this study tested and presented a 10-item long version of the scale, this study views the 5-item version of the scale as likely to be the most useful measurement for most researchers for several reasons. To begin with, the 5-item version of the scale is shorter and thus consumes less space in a survey instrument and imposes less burden on respondents. The fit of the 5-item scale was better than the 10-item scale in our confirmatory factor analysis. Moreover, in our survey analysis, the short and long scales were nearly perfectly correlated ($r = .98$) and had nearly identical correlations with the theoretical criteria of autonomy, budget, and power. In addition, the short scale seems to have somewhat better discriminant validity in terms of its statistical independence from

ideology, age, and income. Thus, for most research and analytical purpose, this study would recommend the use of the 5-item scale. The 10-item scale, however, may be preferable in studies in which the reputation profiles of agencies along the five content domains of bureaucratic reputation are of interest, as that was demonstrated in Figure 2.4 for the FDA, NASA, and IRS. In such analyses, having two items to represent each content domain of bureaucratic reputation would offer additional precision and reliability.

CHAPTER 3:

PREDICTORS OF CITIZENS' REPUTATION JUDGMENTS

From: Lee, D., & Van Ryzin, G. G. (2020). Bureaucratic reputation in the eyes of citizens: an analysis of US federal agencies. *International Review of Administrative Sciences*, 86 (1), 183-200

1. Summary

Bureaucratic reputation has been defined as a set of beliefs about a public organization's capacities, roles, and obligations that are embedded in a network of multiple audiences (Carpenter, 2010). Although one of the most important audiences in a democracy is the citizenry, not much empirical investigation has looked at citizens' beliefs about specific government agencies and what individual or contextual factors influence these beliefs. To examine this question, this study analyzes data from a unique 2013 Pew Political Survey that represents the responses of 1,500 US citizens on the reputations of 12 federal agencies. Results demonstrate that citizens view the reputations of some agencies (such as the CDC and NASA) much more favorably than other agencies (such as the IRS and the Department of Education). In regression analyses, findings suggest that the reputation of federal agencies varies according to citizens' general level of trust in government and their political ideology, but that demographic, socioeconomic and regional differences also shape reputation judgments. These findings provide some preliminary empirical understanding of the reputation of government agencies in the eyes of the citizenry and may have implications for agencies seeking to manage their relationship with the public.

2. Introduction

Bureaucratic reputation has been defined as a set of beliefs about a public organization's capacities, roles, and obligations that are embedded in a network of multiple audiences, including citizens (Carpenter, 2001, 2010). Bureaucratic reputation has received growing scholarly attention in the field of public administration, with a body of studies showing the significant role reputation plays in guiding the behavior of public organizations and shaping their power and effectiveness (Carpenter, 2001, 2010; Maor, 2011; Maor and Sulitzeanu-Kenan, 2013). These studies suggest that government agencies attempt to avoid being viewed negatively by the public and thus engage in strategies that enhance their reputations, such as blame shifting (Gilad, Maor, and Ben-Nun Bloom, 2013). Though researchers have written extensively about the relationship between public organizations and reputation at a macro-level, little work has been done on the formation of reputation judgments by citizens and other audiences at a micro level. Previous studies have tended to rely on proxy data to measure reputation, such as the coverage of agencies in newspapers or websites, assuming that these sources reflect the views of the public and other audiences. However, studying bureaucratic reputation as viewed directly by citizens at the individual level remains important, especially in the context of the US with its large population and a wide variety of cultural backgrounds, ideological views, and geographical settings. It is especially interesting to focus on US federal agencies, as their functions often concentrate on particular segments of the population or regions of the country and thus their reputations may be dependent on associated interests or beliefs (Morgeson and Petrescu, 2011; Meier, 1987).

This study, therefore, provides an exploratory investigation of the reputation of US

federal agencies in the eyes of citizens using data from a 2013 Pew Political Survey. This survey is unique in that, rather than asking about the federal government in general, it taps into citizens' favorability judgments of 12 specifically named federal agencies. Capturing reputation with a survey question has its limitations because public agencies are viewed within a complex network of multiple audiences having pluralistic interests. As Carpenter and Krause (2011: 27) observe, "[c]omplex public organizations are seen 'through a glass but dimly' by their manifold audiences". The Pew survey does a good job, we would argue, of capturing this "dim" view of bureaucratic reputation as seen by citizens. Results demonstrate that citizens view the reputations of some federal agencies (such as the CDC and NASA) much more favorably than other organizations (such as the IRS and the Department of Education (DOE)). In regression analyses, findings suggest that the reputation of federal agencies varies according to citizens' general level of trust in government and their political ideology, but that demographic, socioeconomic and regional differences also shape reputation judgments. The following sections provide the theoretical background, data, and results of our study.

3. Research Background

Reputations are either positive or negative, meaning that they have direction or valence (Walker, 2010). Also, according to Fombrun (2012: 100), reputation can be understood as "a collective assessment" of an organization's "attractiveness" in the eyes of a particular group of stakeholders. Our study relies on this definition in part, as it focuses on the "favorability" of federal agencies in the eyes of citizens. In the field of public administration, bureaucratic reputation has been defined by Carpenter (2010: 45) as "a set

of symbolic beliefs about the unique or separable capacities, roles, and obligations of organizations, where these beliefs are embedded in audience networks.” Indeed, Carpenter’s work has inspired a growing body of research about the importance of reputation in the public sector, initially focused on the relationship between administration and politics (Carpenter, 2001, 2010; Maor, 2011; Roberts, 2006). These scholars argue in particular that favorable reputations give legitimacy to agencies and thus become a source of bureaucratic autonomy and power in a system of democratic governance (Carpenter, 2001).

Another stream of research in this area has focused on the impact of reputation on the behaviors of public agencies. These studies have found that reputation influences bureaucratic performance, administrative choices, and cooperation within networks (Busuioc, 2016; Christensen and Lodge, 2016; Gilad et. al., 2013; Ingold and Leifeld, 2014; Krause and Douglas, 2005; Maor, Gilad, and Bloom, 2012; Maor and Sulitzeanu-Kenan, 2013; Moynihan, 2012). Findings from these studies suggest that public organizations try to avoid being negatively perceived by improving their performance as well as by choosing strategic responses to reduce blame. In addition, findings indicate that agencies attempt to manipulate their reputation as a way to increase their power in policy networks. However, as Moore (2015) showed with the example of the US Department of Veterans Affairs (VA), it is sometimes possible for organizations to have good performance even when they suffer from a poor reputation because of high-profile scandals.

The fact that bureaucratic reputation is characterized by externally formed subjectivity (Carpenter, 2010) implies that factors outside of the actions or performance of a public organization could affect its reputation. These factors include characteristics of the

audience members themselves and the contexts in which they experience or perceive an agency. While reputation is seen as the aggregated perceptions of multiple audiences, it could possibly cover the opinion of particular subgroups of the public (Page and Shapiro, 1992). Citizens have fragmentary and conflicting attitudes based on different values, expectations or interests that can lead them to perceive the same objective circumstances in various ways and thus lead to diverse policy preferences (Chong and Druckman, 2007; Page and Shapiro, 1992). Moreover, citizens' attitudes towards a federal agency tend to be shaped by how they value that agency's functions and responsibilities (Meier, 1987). These considerations imply that even if an agency performs well, some people could still perceive the agency in less favorable terms because of ideological or other attributive factors. Indeed, previous studies have found that general attitudes toward government tend to be associated in predictable ways with various demographic, socioeconomic, geographic and ideological characteristics, which can reflect self-interest as well as cultural and other differences (Page and Shapiro, 1992; Sanders, 1988; Rudolph and Evans, 2005). These relationships, which form the main focus of our empirical analysis, are summarized in Figure 3.1 and discussed in the remainder of this section.

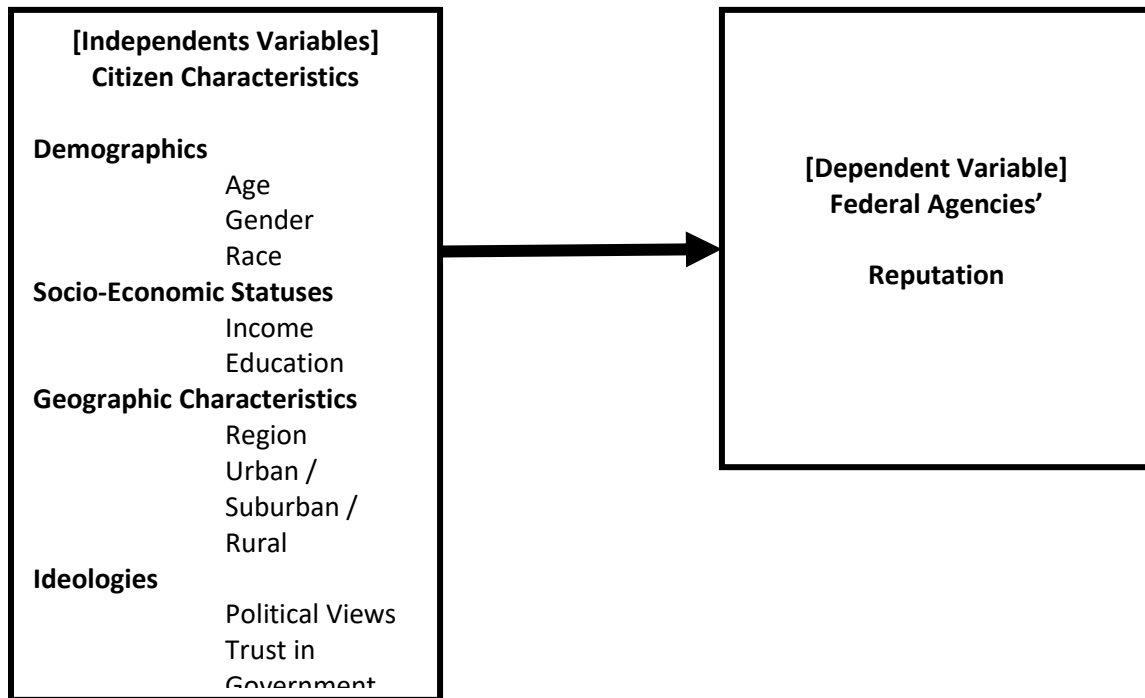


Figure 3.1: Conceptual model of citizen characteristics as determinants of agency reputations

Demographic factors were firstly included, such as gender, race, and age because studies of policy preference and public opinion have found evidence that these factors influence people's perceptions of government in general as well as specific public policy issues. For example, gender has been shown to influence judgments of policy issues such as social welfare, environmental protection, public safety and education (Cook, 1979; Page and Shapiro, 1992; Shapiro and Mahajan, 1986; Rudolph and Evans, 2005). Building on these findings, we might expect that women would be more inclined to view the reputations of related agencies in more favorable terms, such the Department of Health and Human Services (HHS), the Environmental Protection Agency (EPA), the Department of Homeland Security (DHS), and the DOE. Race and ethnicity have also been found to influence policy preferences, particularly for African Americans to be relatively more

supportive of medical care, welfare, educational opportunity, and environment because of their historical experiences of institutional racism and social and economic disadvantage (Page and Shapiro, 1992; Eismeier, 1982). Based on these findings, this study assumes that African Americans may be more likely to perceive agencies like the HHS, DOE, and EPA as having good reputations. The role of age in public opinion has also been emphasized, in part because of cohort effects (historical differences in policy contexts over the lifespan) as well as differences in the needs of older people due to their economic, social, and physical conditions. Many studies have suggested that older people care more about Social Security, Medicare, and law enforcement (Fraile and Ferrer, 2005; Matheson and Wearing, 1999; Page and Shapiro, 1992). Younger citizens, in turn, tend to favor policy areas such as education, the environment, and science and technology (Eismeier, 1982; Sanders, 1988; Rudolph and Evans, 2005). Correspondingly, we might expect older people to view the HHS (which administers Medicare) and the Social Security Administration (although this agency is not included in the Pew survey) as having higher reputations, while younger people might be expected to give more reputational weight to the EPA, DOE and NASA.

The level of socioeconomic (SES), which includes the characteristics of income and education, is also likely to influence reputation judgments. Previously studies have shown that SES is associated with the access to services, social and professional networks, and other resources that shape political attitudes and policy preferences (Bradeley and Corwyn, 2002). Some researchers have suggested that higher SES is associated with support for the arts, human and social rights, and the environment, while lower SES is associated with a preference for social welfare programs, including healthcare, education, and welfare (Busemeyer, Goerres, and Weschle, 2009; Page and Shapiro, 1992). Based on

this prior research, this study expects higher SES citizens to judge the reputations of agencies such as the Department of Justice (DOJ) and EPA in favorable terms, while lower SES people would be more likely view the DOE and HHS as having good reputations.

Geographical characteristics, such as region (the South, Northeast, West, Midwest) and community type (urban, suburban, rural) also could be expected to influence citizens' judgments of agency reputations. One reason for this is that, in the US, federal policies impact regions and communities differently. For example, the West contains an abundance of national forests, parklands, and natural resources. As a result, the citizens of the West might be expected to have a greater interest in the policies of agencies like the EPA. The West is also home to much of the aerospace industry, so it is possible that NASA has a more positive reputation in the region. The Northeast is more urban and densely populated and thus more effected by agencies such as the Department of Transportation and Department of Housing and Urban Development (although neither of these agencies is included in the Pew survey). The South contains many military bases, thus citizens there may be more strongly supportive of defense agencies (Page and Shapiro, 1992; Rice, McLean, and Larsen, 2002). Although there are clear data limitations, and many of these hypotheses remain exploratory, this study nevertheless believes that region, in the US context, has the potential to shape agency reputations as viewed by citizens.

With respect to community type, urban areas often benefit more from federal mass transit and housing programs, so it is likely agencies like DOT and HUD may be seen as more reputable by people living in urban areas. Another factor is the disadvantages of service availability and accessibility in rural areas compared to urban areas (Allard and Cigna, 2008). For example, healthcare service delivery in the US is highly skewed toward

urban areas by the imbalanced distribution of resources and facilities (Ricketts, 2000). Also, there has been a long discussion about urban-rural differences in the impacts of environmental policies, although the results are mixed (Althoff and Greign, 1977; Lowe and Pinhey, 1982). This might suggest that people who live in rural areas would be less likely to prefer the HHS and Centers for Disease Control and Prevention (CDC).

In addition, this study's hypothesis, citizens' views of agency reputations depends a great deal on political ideology, has been referred to as an "underlying principle" of much of the variation in public opinion about government and public policy (Feldman, 1988: 416). For example, previous studies have demonstrated that liberals tend to prefer the involvement of government in solving social problems, including welfare, environmental protection, and education, while conservatives prefer policies that promote less government involvement and regulation, except for national defense (Lavertu and Moynihan, 2013; Rudolph and Evans, 2005). Thus, this study assumes that people who are liberal would form more favorable judgments of the reputations of agencies such as the EPA, HHS, DOE, and CDC, while conservatives would consider the DOD and NSA as having good reputations. Trust of government has been considered as another important measure of political ideology that is often associated with policy preferences. According to sacrifice-based theory, political trust is related to people's support for programs involving widely distributed cost with narrowly concentrated benefits (welfare, health care, public schools) but not programs having universal distribution of benefits (crime prevention, environmental protection, and national defense) (Hetherington, 2005; Rudolph and Evans, 2005). This leads us to expect that people who trust government would be more likely to

assign reputational value to the HHS and DOE, while the reputations of the DOJ, EPA, DHS, NSA and DOD would be less affected by the level of trust.

4. Data and Measurement

As mentioned, this study used the Political Survey conducted by the Pew Research Center in October 2013. Adults 18 years of age or older from all 50 states were interviewed by both landline and cellular telephones (<http://www.people-press.org>). Telephone calls were made by random digit dialing (RDD), a probability sampling method that gives all phone numbers in the U.S. a known chance of being included and thus helps ensure coverage and representativeness. There were 1,500 respondents of which 750 were called on landlines, and 750 were called on cell phones. The final sample size is 1,273, after excluding responses with missing values¹.

4.1. Outcome variable

In organizational theory literature, there are generally two approaches to measuring reputation: an overall measure and an attribute-specific measure (Dowling and Gardberg, 2012). This study follows the former approach and use an overall measure of reputation that represents “a collective assessment” of an organization’s “attractiveness” (Fombrun, 2012: 109). This overall measure is also consistent with some definitions of reputation that emphasize attractiveness, likeability, and admiration (Wæraas and Byrkjeflot, 2012). Thus, this study employs the following question from the Pew survey: “Now I would like to ask

¹ The sample size ranges from 576 to 622 because the survey split 12 agencies into two groups (6 per each) and alternately select one of the two.

you about some parts of the government. Is your overall opinion of [INSERT AGENCY] very favorable, mostly favorable, mostly unfavorable, or very unfavorable?” To facilitate multivariate modeling, as explained more fully in the next section, this study recoded this into a dichotomous variable where 1=very/mostly favorable and 0=very/mostly unfavorable.

In the Pew survey, this overall reputation measure was asked for twelve federal agencies: The Centers for Disease Control and Prevention (CDC), National Aeronautics and Space Administration (NASA), the Internal Revenue Service (IRS), the Department of Education (DOE), the Department of Defense (DOD), the Department of Veterans Affairs (VA), the Food and Drug Administration (FDA), the Environmental Protection Agency (EPA), the Department of Health and Human Services (HHS), the Department of Homeland Security (DHS), the Department of Justice (DOJ), and the National Security Agency (NSA). Other federal agencies are of theoretical interest, as the above discussion suggests, but our empirical analysis of necessity is limited to these 12 agencies. Still, this Pew survey is one of only a few national surveys that has asked about specific federal agencies in such detail.

Figure 3.2 shows the 12 federal agencies in rank order of reputation, from highest to lowest, as measured by the favorability percentage. As can be seen, the reputations of these agencies vary quite a bit, from a high of nearly 90% for the CDC to a low of less than 50% for the IRS. It is interesting that science and health agencies (the CDC, NASA, and FDA) have generally good reputations overall, as do military agencies (the DOD and VA). The IRS is the lowest rated agency, as mentioned, no doubt because of the unpopularity of its main function of tax collection and enforcement. Interestingly, the other relative lower-

reputation agencies, NSA and DOE, have been associated with politically contentious issues in the US, namely surveillance of communications (NSA) and education reform (DOE). More broadly, what Figure 3.2 shows is that federal agencies vary in their overall reputations in the eyes of the public. What individual and contextual characteristics of citizens explain this variation is the aim of our multivariate analysis, which will be explained next.

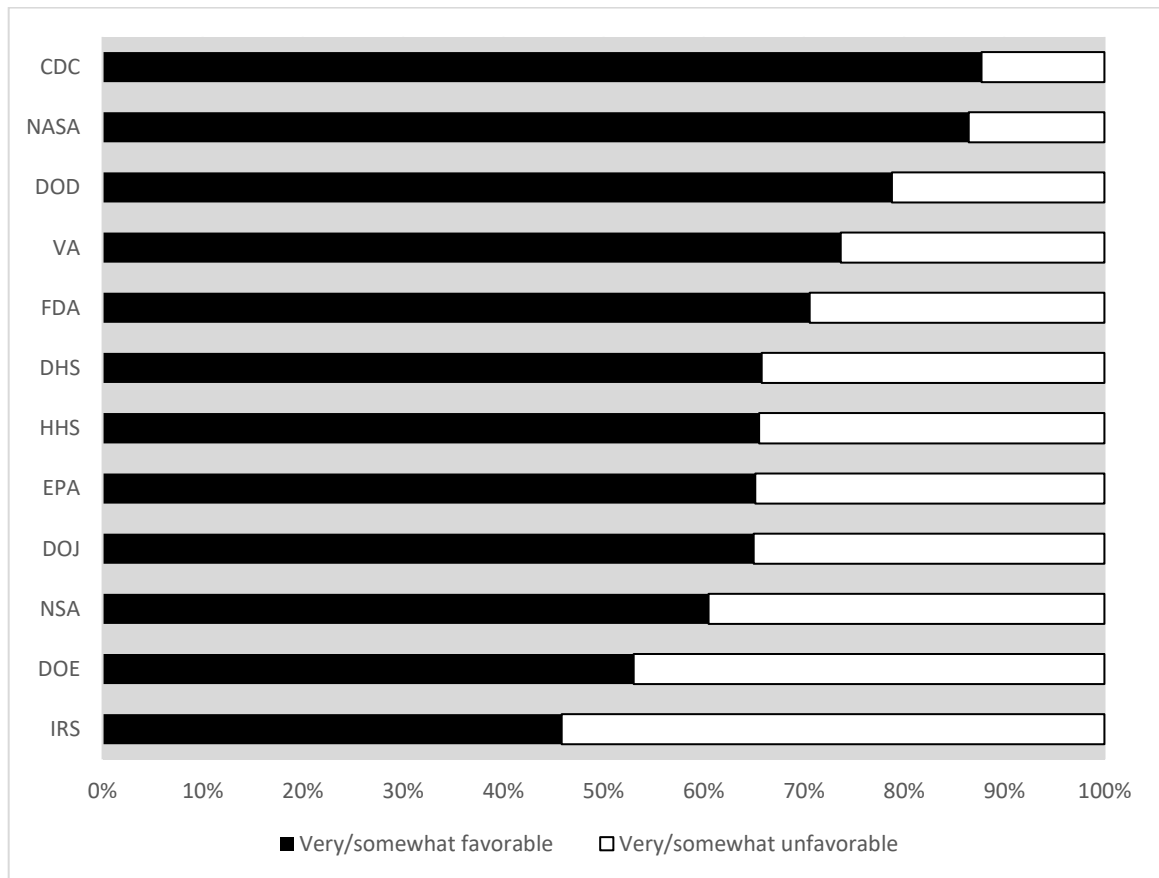


Figure 3.2: Reputations (favorability ratings) of 12 federal agencies (in rank order)

4.2 Predictor variables

Table 3.1 shows the descriptive statistics of the sample in terms of demographics, SES, geographic characteristics, and ideology. The sample is about evenly divided between

females (47%) and males (53%), with an average age of 53 ($SD = 17.4$). Blacks were 10% of the sample, while whites were 79% and Hispanics, Asians and other categories were relatively small. For this reason, we use a dichotomous measure of race (1= black, 0=other). Concerning socioeconomic status (SES), education was scored on an 8-point scale ranging from 1 (less than high school) to 8 (postgraduate or professional degree). Income was measured by a 9-point scale ranging from less than \$10,000 to \$150,000 or more. For geographic characteristics, we include regional dummy variables representing the West, Northeast, South, and Midwest, with the later excluded for purposes of the regression analyses. This study also included community type, with dummy variables representing rural (17% of respondents), urban (33%) or suburban (49%). In the regressions, suburban was excluded as the reference category.

Table 3.1: Question wording and descriptive statistics for independent variables (n=1,273)

Variable	Question Wording	Min	Max	Mean	SD
<i>Demographics</i>					
Gender	Respondent's sex (0=male and 1=female)	0	1	0.47	0.53
Age	Respondent's reported years of age	18	99	52.84	17.37
African American	Dummy variable for respondents who are Black or African-American	0	1	0.10	0.30
<i>Socio-Economic Statuses</i>					
Education	What is the highest level of school you have completed or degree you have received? (1=less than high school, to 8=postgraduate/professional degree)	1	8	5.11	1.79
Income	Last year, that is in 2012, what was your total family income from all sources, before taxes? (1=less than \$10,000, to 9=\$150,000 or more)	1	9	5.38	2.46
<i>Geographic Characteristics</i>					
Northeast	Dummy variable for respondents who live in Northeast	0	1	0.18	0.39
South	Dummy variable for respondents who live in South	0	1	0.36	0.48
West	Dummy variable for respondents who live in West	0	1	0.24	0.41
Midwest	Dummy variable for respondents who live in Midwest	0	1	0.22	0.41
Rural	Dummy variable for respondents who live in rural distinction	0	1	0.17	0.37
Urban	Dummy variable for respondents who live in urban distinction	0	1	0.33	0.47
Suburban	Dummy variable for respondents who live in suburban distinction	0	1	0.49	0.50
<i>Ideologies</i>					
Conservative	Dummy variable for respondents who are conservative	0	1	0.40	0.49
Moderate	Dummy variable for respondents who are moderate	0	1	0.37	0.48
Liberal	Dummy variable for respondents who are liberal	0	1	0.23	0.41
Trust in Govn't	How much of the time do you think you can trust the government in Washington to do what is right? (1=Never, to 4=Just about Always)	1	4	2.08	0.59

Ideology was measured by both political ideology and overall trust in government. Political ideology was measured on a 5-point scale ranging from 1= “very conservative” to 5= “very liberal.” This study recoded this into dummy variables for conservative (40%) and liberal (23%), with moderate (37%) as the excluded category. Also, the survey asked respondents whether they trusted the government in Washington to do what was right, using a 4-point scale ranging from “never” to “just about always.” The mean score of 2.08, which is below the midpoint, suggests that respondents in general had somewhat low levels of trust in government.

Table 3.2: Logit regression analyses of federal agency reputations

	Reputation of federal agency											
	CDC	NASA	DOD	VA	FDA	DHS	HHS	EPA	DOJ	NSA	DOE	IRS
<i>Demographics</i>												
Gender(Female)	-0.08	0.03	0.00	0.03	-0.00	0.18***	0.06	0.15***	0.02	0.10**	-0.07*	0.01
Age	-0.05	-0.09	-0.01	-0.07	0.03	-0.02	0.01	-0.03	-0.05	0.06	-0.02	0.04
Race(Black)	-0.04	-0.06	0.01	0.07	0.10*	0.05	0.11**	0.12**	-0.07	0.08	0.06	0.13**
<i>Socio-Economic statuses</i>												
Education	0.05	0.20**	0.10	0.09	-0.03	-0.13**	-0.11*	-0.05	-0.04	-0.02	-0.09*	0.02
Income	0.22***	0.11	0.13**	0.04	0.15***	-0.02	0.05	-0.03	0.07	0.03	-0.03	0.01
<i>Geographic characteristics</i>												
Northeast	0.14*	0.16**	0.01	-0.03	-0.01	0.11*	0.07	0.03	0.06	0.01	-0.01	0.05
South	0.17**	0.14*	0.02	-0.00	0.03	0.00	-0.02	0.05	-0.01	0.05	-0.03	-0.01
West	0.03	0.13	-0.12*	0.05	0.02	-0.05	0.08	0.07	-0.01	-0.02	-0.11**	-0.03
Urban	-0.10	0.06	-0.05	-0.04	-0.02	-0.09*	0.08	0.04	0.09*	-0.05	0.03	-0.04
Rural	-0.12**	-0.01	0.08	0.03	0.01	-0.08	0.01	-0.06	-0.02	0.01	0.03	-0.05
<i>Ideologies</i>												
Conservative	-0.08	-0.07	-0.02	0.08	-0.15***	-0.13**	-0.25***	-0.26***	-0.17***	-0.12**	-0.16***	-0.30***
Liberal	0.20**	0.09	-0.14**	0.02	-0.01	-0.09*	0.10*	0.13**	0.08	-0.06	0.14***	0.10**
Trust in Govn't	0.15**	0.16**	0.13**	0.19***	0.14***	0.26***	0.32***	0.22***	0.29***	0.22***	0.23***	0.23***
<i>Pseudo r2</i>	0.11	0.10	0.05	0.04	0.04	0.10	0.17	0.17	0.11	0.06	0.10	0.16
<i>N</i>	587	576	622	600	612	614	589	608	611	587	617	620

Note: Table shows logistic regression coefficients; *** p < .01, ** p < .05, * p < .10 (two-tailed tests).

5. Analysis and Results

Because the measures of reputation for each agency were dichotomized, as our dependent variables, binary logistic regressions were employed in our analyses. This study considered maintaining the 4-point ordinal scales and using an ordered logit model but found evidence in our data of violations of the proportional odds assumption. With respect to regression diagnostics, bivariate correlation results show that the variables have only modest intercorrelations (below .5), and tests revealed that there is no problem of multicollinearity. The results of our logistic regression analyses are presented in Table 3.2, which shows logit coefficients and pseudo R-squares. Note that, overall, the models explain a relatively modest proportion of variance in bureaucratic reputation, ranging from a low of 4% (for the VA and FDA) to high of 17% (for the HHS and EPA). Nevertheless, there are some clear patterns of significant predictors of bureaucratic reputation for each of the 12 agencies.

Beginning with demographic factors, the two variables associated most consistently with perceived bureaucratic reputation are gender and race-ethnicity. Specifically, women are more likely to view the DHS and the EPA as having good reputations, as this study expected from prior research. Women also view the reputation of the NSA in more favorable terms, which was unexpected although perhaps similar to the DHS in its focus on security issues. Clearly, contrary to our expectations, women are less likely to rate the DOE as having a good reputation. This can be speculated and may be explained in part by the greater involvement of women in local schools, which in the US have tended to be critical of recent DOE policies promoting more rigorous Common Core Standards and related testing. With respect to race-ethnicity, African Americans perceive the reputations

of the FDA, HHS, and the EPA more positively, which is in line with our assumptions based on prior research. Curiously, African Americans judge the IRS as having a more positive reputation, relative to other race-ethnic groups, which may reflect a greater concern about redistributive policies. Finally, in contrast to our initial assumption, age was not a factor influencing the reputations of federal agencies in the eyes of citizens. This finding is somewhat surprising, given our expectation that cohort influences and age effects would shape the views of federal agency reputations in various ways.

Socioeconomic status appears as a significant predictor of the reputations of federal agencies, although not entirely in line with our expectations. As education level increases, people tend to view NASA more positively, which does fit our expectations. This result could possibly be explained by representative bureaucracy theory, which emphasizes the shared attitudes and values between administrators and certain segments of the public (Meier and Nigro, 1976). As NASA is associated with advances in science and technology, it may be that more educated people identify with the organization and its mission and thus view it more favorably. Somewhat unexpectedly, however, education is negatively related to the perceived reputations of the DHS, HHS and the DOE. Individuals with higher incomes, in turn, are more likely to judge the reputations of the CDC, the DOD, and the FDA in positive terms. Overall, however, income and education are related to the bureaucratic reputation in a somewhat more mixed pattern than that was assumed.

Turning to regional differences, in the Northeast, the CDC, NASA, and the DHS, are agencies that are judged by people as having generally better reputations. In the South, people do not appear significantly distinct from people living in other regions except with respect to their judgments of the reputations of the CDC and NASA (both agencies with a

large presence in the South). In the West, citizens tend to judge the DOE and DOD somewhat less favorably. Taken together, these results confirm the existence of a regional difference in the US in the perceived reputations of various federal agencies, although the pattern is complex and not entirely what was expected. Also, findings suggested some urban-rural differences, with people living in rural areas viewing the reputation of the CDC in a less favorable light, perhaps because of the public health concerns in many rural areas in the US related to chronic disease and (more recently) opioid abuse. In contrast, people in urban areas tend to have a more favorable view of the reputation of the DOJ, while discounting that of the DHS. This may reflect the role of the DOJ in addressing the problems of policing in urban America (at the time of the survey) as well as the general concern with civil rights in more diverse, urban communities. Again, these patterns are complex and not fully in line with our preliminary assumptions.

Ideological factors, including political ideology and trust, emerge as the most consistent and strong predictors of the reputations of federal agencies in our results. As expected, conservatives have unfavorable attitudes towards regulatory agencies like the FDA and the EPA, as well as social welfare agencies like HHS and civil rights agencies like the DOJ. Curiously, conservatives also discount the reputations of security agencies like the DHS and NSA. Liberals, for their part, view the reputation of the CDC positively, as well as HHS, the EPA, and the DOE. In line with our expectation, liberals tend to have negative views of security agencies like the DOD and the DHS. The IRS is an especially interesting agency with respect to ideology, as it is viewed quite negatively by conservatives yet positively by liberals, most likely because of deeply divided views of the appropriate size and role of government. Finally, generalized trust in government is

positively and significantly associated with all federal agencies' reputations, meaning that people with higher trust gave higher reputation scores for every federal agency.

6. Conclusions

Somewhat expectedly, the reputations of federal agencies are strongly associated with how much citizens trust in government generally. Those who trust government more judge the reputations of most agencies better than people who trust government less. This is in line with previous studies pointing out the relationship between the image of government and trust in government, which find that negative images of government organizations stem from the decline of general trust in government (Waeraas and Byrkjeflot, 2012). These findings have some implications for the debate about attitudes towards the public sector, specifically whether it is influenced by performance or more deep-seated stereotypes (Van de Walle, Van Roosbroek, and Bouckaert, 2008). What can be inferred is that regardless of the performance of the 12 federal agencies in this study, citizens' judgments appear to be strongly influenced by their pre-existing trust of government in general.

Also, this study found that political ideology is an important factor, suggesting it may serve as a heuristic that citizens use to make inferences about a public organization, as Levertu and Moynihan (2013) suggest. For liberals, findings suggested that contrasting attitudes towards so-called 'liberal agencies' and 'conservative agencies.' Specifically, liberals give higher reputation ratings to agencies that reflect liberal policy priorities, such as the environment, education or health and human services. In contrast, liberals tend to view agencies having responsibility for national defense less favorably. Contrary to our

expectation, conservatives view both liberal and conservative agencies less favorably, perhaps a reflection of their overall anti-government posture. (Goodsell, 2014). Because the Pew survey was conducted in 2013 when Barack Obama was president and Democrats controlled the Senate, the political moment of the survey also likely influenced how conservatives rated federal agencies. It would be interesting to study how these views might have changed under President Trump and a Republican-controlled Congress. Interestingly, conservative ideology is only associated with the less favorably viewed federal agencies but not the four most highly-rated agencies. This may be explained by that the agencies with lower reputations tend to be more vulnerable to political ideology, and especially prone to criticism by conservatives.

Many of the other potential predictors, including demographics, SES, and geographic characteristics have relatively weak relationships with bureaucratic reputations. Although previous studies suggest age influences policy preferences, our results show that age was not a strong predictor of agency reputations. Also contrary to the policy preferences literature, findings did not suggest a link between SES and the reputations of civil rights or environmental agencies. However, findings did provide an association between SES and the reputation of agencies involved in health, education, and security. Taken together, these results suggest that traditional predictors of policy preferences do not completely explain bureaucratic reputations.

Also, this study found some interesting inconsistencies in reputations between parent agencies and their subsidiaries. Our results show that the determinants of the reputation of the HHS are different from those of the CDC and FDA, which is somewhat surprising considering that the HHS is the parent agency of the CDC and FDA. The NSA,

which is a part of the DOD, likewise shows a different pattern of determinants than its parent agency. These results indicate the possibility that each federal agency has its own somewhat independent reputation, and thus agencies within a larger department can and do appeal to different audiences in the large society. More research should be done on this topic, including studies that look at the implications for bureaucratic politics and management when different agencies within the same department appeal to different audiences to bolster their reputations.

This study has some limitations as well as implications that future research that seeks to understand how citizens view bureaucratic reputation. To begin with, questions on favorability were used as proxies for the reputations of federal agencies. Favorability ratings reveal the overall perceptions of an agency, but “favorability” may not encompass all of the various dimensions that make up an agency’s reputation, such as performance, morality, procedural fairness, and technical competence (Carpenter 2010). More effort should be made to improve the measurement of bureaucratic reputation in the eyes of citizens, including the development and validation of multi-item scales with better reliability as well as content validity along these key dimensions.

This study’s limitation should be noted that our assessment of construct validity may be influenced by common source bias, particularly to the extent respondents have generally positive or negative feelings about a given agency that may underlie their responses to questions about reputation, budget, power, and autonomy. But still our scale also demonstrated discriminant validity (lack of correlation) with other survey variables and corresponded fairly closely to the findings from independent sources such as the Pew survey.

CHAPTER 4:

REPUTATION MANAGEMENT

AND COMMUNICATION STRATEGIES

1. Summary

Scholarly attention has focused primarily on blame avoidance strategies, particularly communication strategies, as a strategic behavior that public organizations may engage in when responding to reputational threats. Nevertheless, the question of those strategies' efficacy in protecting an organizations' reputation still remains understudied. Neglect of this research question restricts our understanding of public organizations' management of their external audiences from which they draw important resources, including bureaucratic reputation. Based on attribution theory and cognitive dissonance theory, this article probed various assumptions about public organizations' communicative blame avoidance responses and tested their effects on citizens' reputation judgments. The results from survey experiments showed that the way an agency explains its failure is important in shaping its bureaucratic reputation, which sheds light on the importance of communication in restoring and protecting an organization's reputation.

2. Introduction

Managing external constituencies is one of the critical functions of public management, as the press and public are key audiences of public organizations (Allison, 1982). The support or opposition from external environments, i.e., the distribution of sentiment, is a fundamental source of public organizations' power and legitimacy, which is not merely conferred by external actors but is a factor that organizations can influence

by their actions (Wamsley & Zald, 1973). Bureaucratic reputation, which relies on external audiences' perceptions of public organizations, becomes a valuable political asset that can be used to generate public support when it is managed successfully (Carpenter, 2002, 2010). When public organizations encounter failures in their operation, using blame avoidance strategies—particularly strategic ways to manage communication—becomes essential, as it protects organizational reputations from consequential threats (Hood, 2012; Moynihan 2012). Scandals, poor service quality, and failure to achieve policy goals change external audiences' judgments adversely and abruptly; the appropriate use of strategies can help maintain and restore reputations effectively.

This study focused primarily on the presentational strategies that Hood (2003, 2012) suggested, which refer to an organization's attempts to affect public attitudes and perceptions to avoid blame by selecting specific arguments. In order to manage public attitudes after a failure is revealed, public organizations need to respond attentively to audiences' perceptions; thus, their chosen presentational strategies are highly likely to affect the agency's reputation (Hood, 2012; Maor, Gilad, & Bloom, 2012). Most studies in the public administration literature have focused on agency and policy strategies. Another two strategies Hood (2012) suggested emphasize formal and institutional settings to prevent blame (Hood 2003, 2012; James, Jilke, Petersen, & Van de Walle, 2015; Marvel & Girth, 2015; McGraw, 1990; Nielsen & Baekgaard, 2015; Sievert, Vogel, Reinders & Ahmed, 2019; Steffel, Williams, & Perrmann-Graham, 2016). Thus, this view suggests it is important to study presentational strategies, although far fewer studies have focused on the subject compared to those on policy and agency strategies.

The bureaucratic reputation literature includes a branch of reputation management research that focuses on the strategic use of communication, which also shares ground with the logic of presentational strategies. Previous research has shown an interactive nature in the relations between organizations and the public, indicating that organizations are less likely to respond to external audiences in an ad hoc way but in a more calculated way (Maor, 2015). Despite the importance of this interactive nature, those studies that have focused on presentational strategies largely have adopted the organizational-level approach that examines the organizations' selection of a strategic response in response to their reputational concern (Gilad, Maor, & Bloom, 2013; Maor, et al., 2012; Moynihan, 2012). Little individual-level research has been reported on the way the external audiences' reputation judgments change in response to agencies' strategies.

Thus, the purpose of this study was to examine the effect of presentational strategies on external audiences, particularly citizens' reputation judgments after an agency failure occurs. According to attribution theory, types of failure and messages play a key role in shaping audiences' causal perceptions and to whom they assign responsibility for failures, which ultimately influences their judgments of organizations (Weiner, 1985, 2000). Based on Coombs and Holladay's (1996) testable assumptions, this study looks at whether the efficacy of a strategy aimed at influencing citizens' reputation judgments depends on the types of failure. Further, based on cognitive dissonance theory (Kunda, 1990), this study probed the role of people's prior judgments on the efficacy of presentational strategies; specifically, it examines whether people tend to be influenced more by information that supports their prior judgments.

This study tested these proposed hypotheses using data from a randomized survey experiment in which the treatments in the vignettes derived from actual media coverage of federal agencies. While previous studies that focus on presentational strategies have analyzed media coverage as a data source (Hood, 2003; Mortensen, 2012), they have had certain limitations in explaining the causal effects of these strategies on external constituents' perceptions or judgments. Taking advantage of within- and between-subjects experimental designs to increase statistical power, this study tested the effect of blame avoidance strategies on citizens' reputation judgments and their interaction with citizens' prior judgments.

3. Research Background

3.1 Blame avoidance strategies and bureaucratic reputation

Blame is central to accountability in the public system, as accountability implies "...having to bear the blame" (Bovens, 2007, p. 189; Romzek & Ingraham, 2000). Blame plays a significant role in holding policy actors accountable, as it threatens actors' important resources, such as reputations. Bureaucratic reputation—a set of public organizations' beliefs that external audiences share—is an essential political asset of public agencies related to their autonomy, power, and legitimacy (Carpenter, 2010). Because they exist and operate within the external environment, public organizations have substantial concerns about citizens' attributions of responsibility for failure, because citizens, who play a role as voters, evaluators, and an ultimate principal in governments, are those who assign blame ultimately in the blame game (Sulitzeanu-Kenan & Zohlnhofer, 2019).

As blame for a failure includes preventable and remediable effects, the primary objective of blame avoidance strategies is to maintain and protect organizations' reputations (Hood, 2012; Moynihan, 2012). In the bureaucratic reputation literature, empirical evidence has supported that concern about protecting reputation is a strong motivator to use different blame avoidance strategies, in which Maor (2015) categorized the research into two streams: the first is the strategic use of communication, and the second is a selective response through changes in decision timing, decision observability, and agency output (Carpenter, 2002; Maor & Sulitzeanu-Kenan, 2013, 2016). While the second is studied more popularly in the current literature, the first is vital, given that presentational strategies are an active way for public agencies to mitigate negative public attitudes to failures (Coombs & Holladay, 1996; Hood, 2012; Maor et al., 2012). Nevertheless, while previous studies have examined the way public agencies calibrate their presentational strategies in response to external judgments of their reputations (Gilad et al., 2013; Maor & Sulitzeanu-Kenan, 2013), less attention has been paid to the way individual audiences, primarily citizens, react to organizations' behavior.

A presentational strategy refers to an argument that organizations select to minimize or avoid blame by using "spin," timing, stage management, and various forms of persuasion to attempt to affect public attitudes or perceptions (Hood, 2003, 2012). While the terms communication and presentational strategy are used interchangeably, this study uses Hood's (2012) concept of *presentational strategy* for consistency. Scholars have proposed multiple types of strategies (Gilad et al., 2013; McGraw, 1990; Hood, 2012) that can be categorized into the following: *denying response*, *admitting response*, and *strategic silence*. The first basic strategic option that public organizations have is the *denying*

response, which indicates that officeholders maintain that there is no problem, loss, risk, or harm so that blame can no longer arise (McGraw, 1990, 1991). Organizations can also deny their responsibility for problems in the form of excuses—admitting the problem, but saying they are not the cause of the incident—or justification—admitting their responsibility, but suggesting that the outcome is desirable (Gilad et al., 2013; Hood, 2012; McGraw, 1990), and this approach is what this study adopted as the definition of a denying response. The second option is an *admitting response*, in which the organization apologizes in the blame process to demonstrate its honesty and sincerity, and that it will do things better in the future, which can bring an end to the problem (Hood, 2012). The third strategy used to protect an agency's reputation is *strategic silence*, referring to "...silence as passivity and relinquishment" (Brummett, 1980, p. 293). Strategic silence is often consistent with maintaining a low profile and secrecy, and refusing to respond in any way (Hood, 2012).

Understanding presentational strategies and citizens' reactions is not intended to justify disguises, schemes, or superficial opportunism, but to improve accountability and responsiveness. First, presentational strategies help public organizations enhance the public and media debate over the issues by providing a "...counterweight to negativity bias" that could have remained unchallenged (Hood, 2012, p. 165). Second, the use of presentational strategies may lead to subsequent increased accountability by providing the public with more and better information in future. Responses from government organizations themselves can be useful information for citizens (Hood, 2012). Third, these presentational strategies are also crucial in government accountability, as one of blame's functions is that it brings "public catharsis" (Bovens, 2007; Moynihan, 2012). Lastly,

public organizations' response to the public itself is their duty and involves their accountability to them.

Despite the importance of studying the effect of presentational strategies in the public administration literature, the number of empirical studies remains insufficient for the following reasons: firstly, there have been no relevant theoretical frameworks that provide grounds for testable assumptions; secondly, there have been methodological limitations to test the causality of the strategies; and, lastly, there has been no relevant measure to capture reputation judgments at the individual level. To fill the gaps in the literature, the next section provides hypotheses and models based on Coombs and Holladay's (1996) symbolic approach, and an explanation of the experimental methods and Lee and Van Ryzin's (2019) Bureaucratic Reputation Scale.

Coombs and Holladay's (1996) work bridges the discussion of blame avoidance and bureaucratic reputation by suggesting the crisis management framework in the communication literature. According to their approach, crisis response strategies are important to the protection of an organization's reputation (image) by mitigating public blame for a crisis. Building upon their testable models, this study developed hypotheses that test the association between communication strategies and reputation. In this dissertation, Coombs and Holladay's (1996) model is applied to the discussion of blame avoidance and bureaucratic reputation, which will be explained in depth in next section.

3.2 Hypotheses and models

Attribution theory refers to the perception or inference of causality and explains how individuals' attribute an event to various causes (Kelley & Michela, 1980). According

to the theory, Weiner (1985) suggests three causality dimensions that people use to form attributions: stability, locus (intentionality), and controllability. Stability refers to whether the event occurs frequently (stable) or infrequently (unstable); locus refers to whether the cause is intentional or unintentional, and controllability refers to whether or not the event's cause is controllable (Weiner, 1985; Wilson, Cruz, Marshall, & Rao, 1993; McAuley, Duncan, & Russell, 1992). Causal attribution based on those dimensions is known to be a determinant of individuals' affective reactions to success and failure (Weiner, 1980), and has been applied widely as a determining factor in the fields of education, law, and health.

Based on these three causal dimensions Weiner (1980) proposed, failure situations are categorized according to two dimensions: intentionality (unintentional vs. intentional), and controllability (external vs. internal). According to these dimensions, he suggested four types of failure: faux pax, terrorism, accident, and transgression. Further, by adding stability (a repeated vs. one-time occurrence) as a third criterion, there are eight proposed situations. Among those, this study focused on one-time accidental and repeated transgression failures, both of which events are controllable internally, but differ in their intentionality and stability. Single accidents are situations that occur only once during normal organizational operations and are attributable to human errors, such as workplace injuries or product defects, while transgressions are repeated situations that involve intentional actions on the part of an organization.

Theoretically, different types of failure influence people's attributions differently because their judgments of those three causal dimensions shape their feelings and behaviors toward the actor. People's evaluations of organizations' responsibility for the failure are determined by its cause (Coombs & Holladay, 1996; Weiner, 1985; Weiner,

Perry, & Magnusson, 1988; Wilson et al., 1993). When people attribute responsibility for the failure, they perceive that organizations are more responsible for repeated events than for a single one (Griffin, 1994; Weiner, 2000). In addition, when intentionality is high, indicating that the failure is less justifiable, organizations' responsibility for the failure is perceived to be stronger than when intentionality is low (Coombs & Holladay, 1996). Thus, compared to a single accident, repeated transgressions lead the public to perceive that organizational responsibility for the failure is more serious, as an organization committed it intentionally.

Blame Hypothesis: Under the condition of a transgression, citizens judge organizations' reputations more severely than under the condition of an accident.

According to Weiner et al. (1988), messages shape perceptions of three causal dimensions and, in turn, the feelings the attribution creates. Coombs and Holladay (1996) provided a valuable framework that explains the way communication is used in attempts to affect peoples' attribution and the subsequent reaction attached to those attributions. This framework, referred to as the symbolic approach, shares an assumption that the organization's response strategy shapes how people's attributions are influenced by the three dimensions. Accordingly, people evaluate an organization's responsibility for a failure depending on their perception of the failure's cause: "The more publics attribute organizations' responsibility for the failure, the stronger the likelihood is of public developing and acting upon negative images of the organization" (Coombs & Holladay, 1996, p. 282).

Coombs and Holladay (1996) provided hypothetical models of organizations' responses that are more effective in certain failure situations in reducing organizations' reputational damage by mitigating adverse public attitudes. With respect to a one time

accident, where situations are unintentional or justifiable, such that organizations seem to have less responsibility, the model suggests that denying responsibility will be more effective, arguing that there was no failure or blame because the organization was not responsible for the failure. In comparison, in the case of repeated transgressions, where the failure involves an organization's repeated and intentional behavior, admitting responsibility and taking action to overcome the crisis will have a more positive effect. Thus, this study assumed that an *admitting response* is a more effective strategy in the case of repeated and intentional failures, while a *denying response* is more effective when the failure occurs once and is unintentional.

Avoidance Hypothesis (a): *When the failure is a onetime accident, the denial response has a more positive effect on citizens' reputational judgments.*

Avoidance Hypothesis (b): *When the failure is a repeated transgression, the admitting response has a more positive effect on citizens' reputational judgments.*

Table 4.1: Classification of failure types and matched response

		Stability	
		Onetime	Repeated
Intentionality	Intentional		Transgression – Admitting strategy
	Unintentional	Accident – Denying strategy	

In addition, when citizens respond to negative information, their prior beliefs about the agency can play a critical role. Festinger's (1957) cognitive dissonance theory explains individuals' tendency to pursue opinion-reinforcing information over opinion-challenging information to reduce dissonance, and it is possible that psychological mechanisms lead individuals to disconfirm negative information and confirm positive information about an organization in which they are in favor. While a series of researchers have investigated the benefits of a favorable reputation, this study focuses on the *halo effect*, which indicates that

a favorable reputation can protect an organization from reputation loss (Caruana, 1997; Coombs & Holladay, 2006). One of the key features of a good reputation is "...protecting the public agency from political attack" (Carpenter, 2002, p. 491). Further, a specific failure is less likely to influence the public's beliefs about organizations, as people believe the agency behaves in a trustworthy way in general (Christensen & Lægreid, 2005). Many studies have shown that people who judge an object favorably are influenced less by negative information because they reject or misinterpret it, given that individuals' cognitive processes are driven by the desire to reduce the discomfort of dissonance (Baekgaard, Christensen, Dahlmann, Mathiasen, & Petersen, 2019; James & Van Ryzin, 2017; Jilke, Van Ryzin, & Van de Walle, 2015; Kunda, 1990). Thus, this study predicted that an agency's favorable reputation can soften the effect of a reputational threat.

In particular, this suggests an interaction effect between citizens' prior beliefs and the organization's strategies. Hood (2012) argues that in addition to time and place, attitudes and beliefs are important factors that influence the effectiveness of strategies of blame avoidance. When organizations provide an explanation for their alleged failure, individual constituents can either accept or reject the account as a response (McGraw, 1991). In the process of making these decisions, people who judge an organization more highly generate a different interpretation of its behavior, judging that risks are low and benefits high, which produces different reputational consequences (Finucane, Alhakami, Slovic, & Johnson, 2000). Previous studies in the marketing literature have found that consumers seek and weigh an account that confirms their prior expectations of an organization selectively (Darley & Gross, 1983). If people have a high expectation of the organization, they are more likely to believe an account of the incident that confirms their

prior beliefs compared to one for which they have a lower expectation (Dawar & Pillutla, 2000). Thus, this study predicted that a blame avoidance strategy, particularly a denying response, is more effective when people judge the organization's reputation more positively, as the information confirms people's beliefs.

Bias Hypothesis: People who already have more positive reputation judgments are influenced less by organizations' failure when organizations use a problem denial response

4. Research Design

Experimental design

This study used a vignette survey experimental design to assess the effect of presentational strategies on citizens' reputation judgments of four U.S. federal agencies (Department of Veterans Affairs (VA), Internal Revenue Service (IRS), Centers for Disease Control and Prevention (CDC), and US Postal Services (USPS). Hood (2012) expressed concern about insufficient social science techniques to assess the effect of institutions' strategies, and Oliver et al (2016) have suggested experiments as a relevant tool for such an assessment. Importantly, the randomization in a survey experiment enables a researcher to estimate the causal effects of blame avoidance strategies on citizens' reputation judgments by randomly allocating different strategies to treatment groups of participants. Further, the vignette design is pertinent in this study, as it is common for public agencies to communicate through written media that influence citizens' reputation judgments (Hood, 2012; Maor, 2016). The survey experiment was managed with Qualtrics software, through which the subjects received the URL to an online questionnaire.

Figure 1 shows the study's experimental design. The experiments employed both a within and between subject design that included four different treatments. The within-

subject design is relevant to test this study's hypotheses that aims to capture the change in an outcome, allowing removal of subject-to-subject variation to increase power. However, this design is subject to confounding factors, such as carry-over effects, when an experience of the pre-measure may influence the post-measure. As a way to prevent this problem, this study had respondents answer a number of demographic and attitudinal questions in between the pre- and post-measure to 'wash out' the effect of pre-measure (Seltman, 2014). Respondents' prior judgments were measured by answering items extracted from the Bureaucratic Reputation Scale (Lee & Van Ryzin, 2019), which captures people's reputation judgments of public organizations. The responses were the aggregated scores of two items measured on a 5-point Likert scale that was standardized later in the analysis. The items in the original scale have high internal validity (>0.9) when the five items are aggregated.

The respondents were asked to read a news article that contained negative information about two of the four federal agencies listed above. To reduce the respondents' fatigue, only two of the four agencies were selected randomly and assigned to them. The selection of agencies and the news articles were based on the attribution theory's three criteria of a crisis situation—locus, stability, and controllability (Wilson et al., 1993). Because stability refers to an organization's history of repeating a similar event, the IRS and VA were selected for their nation-wide scandals (Meier & Bothe, 2007; Moore, 2015). In comparison, the USPS and CDC were selected for their consistent high reputations among the public. To establish the validity of their reputations, the Political Survey data from the Pew Research Center from 2010-2017 were checked and confirmed that the favorability score of the IRS and VA dropped after the scandals, while those of the USPS

and the CDC were highly stable for seven years. This also supported previous researchers' argument that articulated the relations between the two agencies' scandals and public attitudes (Moore, 2015; Meier & Bothe, 2007)

The study's vignette design is shown in the Appendix, which contains negative statements about public agencies in the form of news articles. These treatment scenarios derived from various actual news sources, including AP News, CNN, and *USA Today*. The news sources were removed to prevent bias attributable to people's attitudes toward such sources. Changes were made in the articles only for experimental purposes. The selection of the negative events in the news articles followed the three dimensions of Attribution theory. The types of failure varied depending on agencies; the articles selected were similar in controllability, but differed in intentionality and stability (Wilson et al., 1993). In particular, the cases for the IRS and VA were considered repeated transgression situations, as both have a history of nation-wide scandals, such as targeting scrutiny and poor services to veterans, respectively. Compared to the two cases, the USPS and CDC's cases involved more single accidents with no history of the same event and less intentionality involved at organizational levels. To confirm validity of the treatments, a manipulation check was conducted in May 2019 (Table 4.3), which in general, supported the distinction between the crisis types and performance history. This will be discussed in more detail in the Results section.

The experimental factor was the four presentational strategies: the deny response, admit response, strategic silence, and no action. No action was assigned to a control group. After the subjects read the vignette, they were assigned to one of the four treatment groups that provided a description of the agency's presentational strategies: 1) Admitting response;

2) Denying response; 3) Strategic silence, and 4) No action. Presentational strategies involve public organizations' efforts to avoid blame by "spin," state management, and argument (Hood, 2012), and the reputation literature has categorized the strategies of strategic silence and admit and deny broadly (Gilad et al., 2013; Maor et al., 2012). The difference between strategic silence and no action is as follows: Strategic silence is the case in which the agencies do not respond to the problem, and 2) No action is the case in which no statement at all is provided to explain what the agency did. This study included no action as a baseline. The respondents received one of the four treatments randomly. The only difference in the treatment was the short sentences that described the agency's different strategic responses. Thereafter, a question that asked about the agency's reputation was included as a dependent variable measured with the Bureaucratic Reputation Scale (BRS) using a 5-point Likert scale (1=strongly agree to 5=strongly disagree).

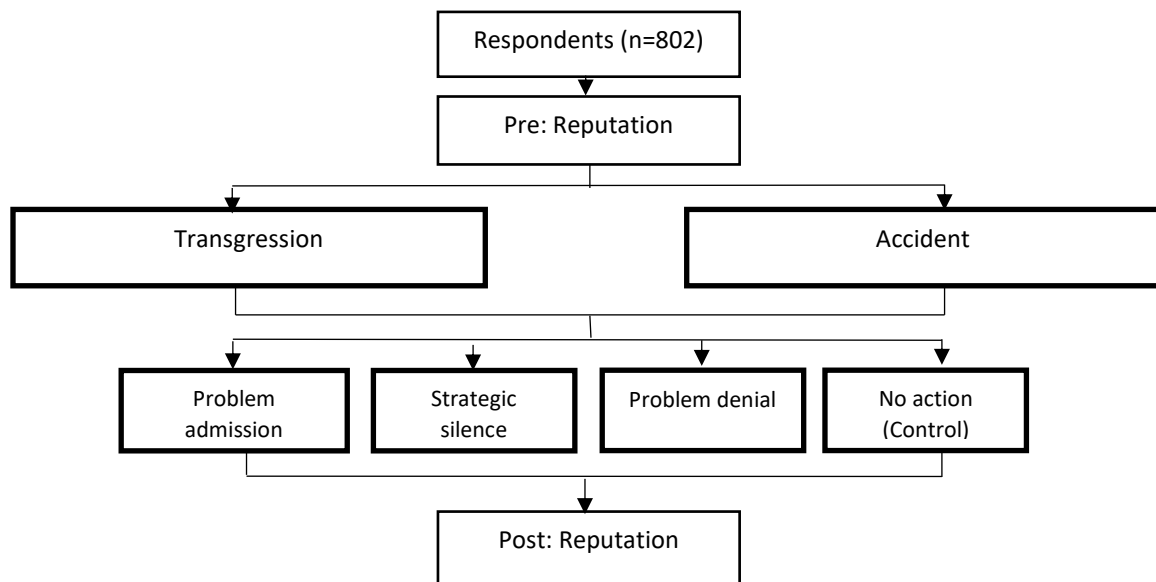


Figure 4.1: An experimental flow chart

Participants

The study's survey experiment included 802 U.S. citizens over 18 years old from all 50 states. The recruitment was conducted in August 2019 and the subjects were recruited from an online panel operated by Qualtrics, a private company that provides online panels. It recruits samples from over 20 market research panels and provides samples based upon client-base requests (Qualtrics, 2014). For this experiment, Qualtrics recruited 1,050 panel members to participate in the survey. While online panels may have the disadvantage of sampling bias, Qualtrics panel partners select respondents randomly for surveys where respondents are highly likely to qualify. Further, upon a request to recruit samples that represent the U.S. population, Qualtrics uses four parameters—gender, race, age, and region—and screens the panels to meet their proportions in the U.S. population. The descriptive demographics for our sample are shown in Table 2. After we excluded the respondents who did not pass the attention check, 802 remained.

Table 4.2: Final sample descriptive statistics

	Freq.	%
Gender		
Female	429	53.49
Male	373	46.51
Race		
White/Caucasian	517	64.46
Black/African American	80	9.98
Hispanic/Latino	136	16.96
Asian	46	5.74
Other	23	2.87
Age		
18-24	82	11.80
25-34	119	19.08
35-44	124	18.05
45-54	142	16.00
55-64	154	16.42
65 or more	181	18.65
Region		
Northeast	124	15.46
South	312	38.90
Midwest	187	23.32
West	179	22.32
Education		
High school or less	189	23.98
2-year/4-year College degree	490	62.18
Masters/Doctoral/professional	109	13.83
Income		
Less than \$34,999	304	38.58
\$35,000- \$74,999	285	36.17
\$75,000 or more	199	25.25
Partisan		
Democrats	291	36.98
Independent	273	34.69
Republicans	223	28.34
N=788-802		

Manipulation check

To check the experimental manipulation, a pilot study that recruited subjects via Amazon's Mechanical Turk (n=262) was conducted in December 2018. Table 3 presents

a series of results of a *t*-test to assess the criteria, including intentionality and stability, used to categorize the failure as either a transgression or accident. After receiving the treatment scenario for the IRS, VA, CDC, and USPS, the respondents answered questions pertaining to their opinions about the intentions of the agencies involved in the event and whether the event was repetitive. The results of the *t*-test showed that the respondents perceived that transgressions involve significantly higher intentionality (0.6 vs. 0.24) and greater stability (repeated occurrence) (0.72 vs. 0.56) than do accidents, which supported the expectations of this research.

Table 4.3: T-test of manipulation check of the scenario

	Transgression	Accident	Difference
Stability	1.72	1.56	0.15***
Intention	1.59	1.24	-0.36***
N= 130			

5. Analysis and Results

This study's analytic approach to test the efficacy of presentational strategies on citizens' reputation judgments was as follows. First, to test the Blame hypothesis, we compared the difference between the mean score of the pre- and post-reputation judgments by failure types. Second, OLS regression was performed to test the Avoidance hypothesis and Bias hypothesis. This analytic approach is appropriate either to control the effect of pre-reputation judgments (Avoidance hypothesis) or include in the interaction model (Bias hypothesis). Third, a comparison among agencies was conducted to illustrate the validity of the study.

Blame Hypothesis

First, the study included an initial analysis of the effect of failure type on the change between pre- and post-reputation judgments. Figure 2 shows the comparison of the means between pre- and post-reputation scores by failure types. As we can see, the respondents who received the transgression vignette reduced their reputational scores from 2.85 to 2.40, and those who received the accident vignette reduced them from 3.62 to 2.93. The bar graph shows that there was a greater reputational loss in the case of accidents than transgressions. In the case of transgressions, it decreased 15%, while in the case of an accident, it decreased 19%. Thus, there was a clear decrease in the change of reputation judgments depending upon the type of failure. However, the results were the converse of the study's expectation that respondents' reputational judgments would decrease more in the case of transgressions than of an accident.

Further, a *t*-test on the difference between pre- and post-reputation judgments indicated that, compared to people's pre-reputation judgments, their post-reputation judgments were significantly worse ($M=2.84$, $SD=0.97$) after they received the transgression vignette ($M=2.40$, $SD=0.95$); in the case of the accident vignette, respondents' post-reputation judgments ($M=3.63$, $SD=0.83$) also were significantly worse than their pre-reputation judgments ($M=2.93$, $SD=0.99$). This result indicates that the treatment was effective in altering people's judgments for the worse. Further, a *t*-test was conducted on the differences between the reputational loss for transgressions (0.44) and accidents (0.69) in an additional analysis, the difference was significant ($t=-6$, $p<0.001$). The reputation judgments of respondents who received the accident vignette decreased more than did those who received the transgression vignette.

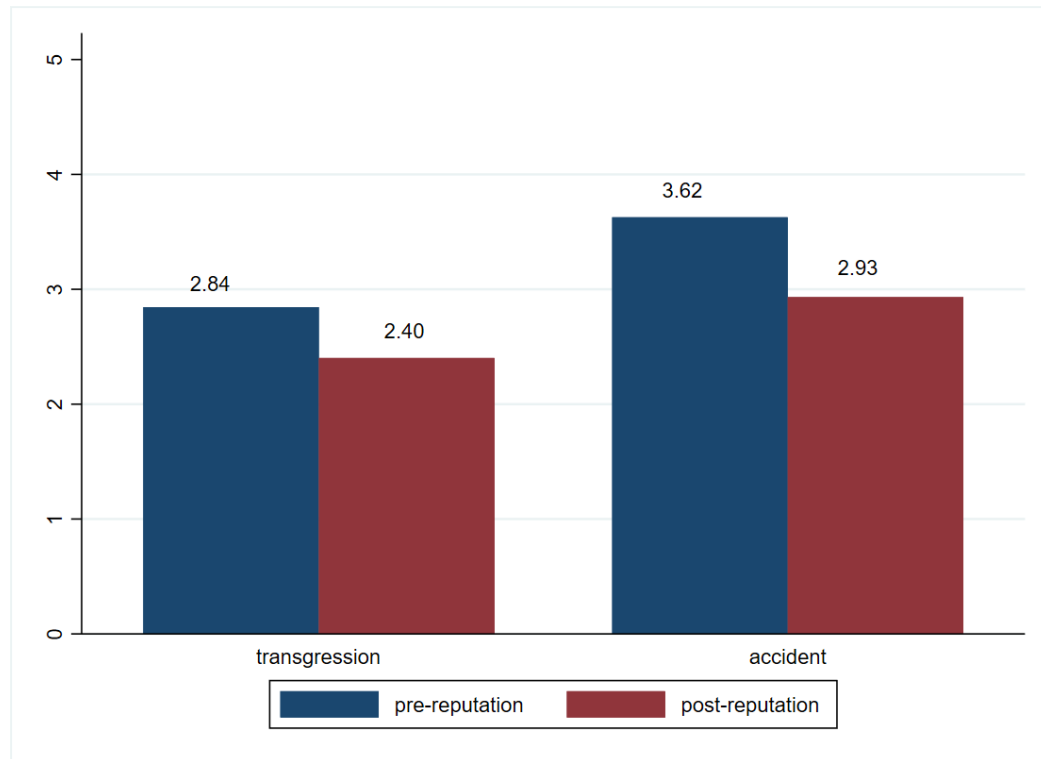


Figure 4.2. Mean comparison of pre and post reputation score by failure types.

Avoidance Hypothesis

The next examination was designed to test whether certain organizations' responses to certain types of failure mitigate the adverse effects on citizens' reputation judgments more effectively. Figure 3 shows the difference between the pre- and post-reputation judgments to the strategic responses sub-grouped by types of failure. The bar graph indicates that for transgressions, the respondents' pre- and post-reputation judgments changed the least for the denying response (-0.35) compared to the other responses: admitting response (-0.46), strategic silence (-0.44), and no response (-0.53). This result was also the opposite of the Avoidance hypothesis (a)'s prediction, that the admitting response would be more effective in the case of transgressions. In addition, the change in

judgments to no response was the greatest among the four, which indicates that explaining is important in mitigating people's blame. Further, in the case of an accident, the denying response led to the smallest difference between the pre- and post-reputation scores compared to the admitting response (-0.67), strategic silence (-0.88), and no response (-0.73). These results supported Avoidance hypothesis (b), that the denying response was the most effective response in the case of an accident.

Next, regression analyses were conducted to test the significance of these differences. The estimated models in Table 4 include the experimental treatments of the four types of strategic responses separately (admitting response, denying response, strategic silence, and no response) as independent variables by types of failure (transgression and accident). Further, the pre-judgment was controlled in the models, and we used a standardized reputation score, which provides means to compare scores that differ from a normal distribution (Cardinali, 2018).

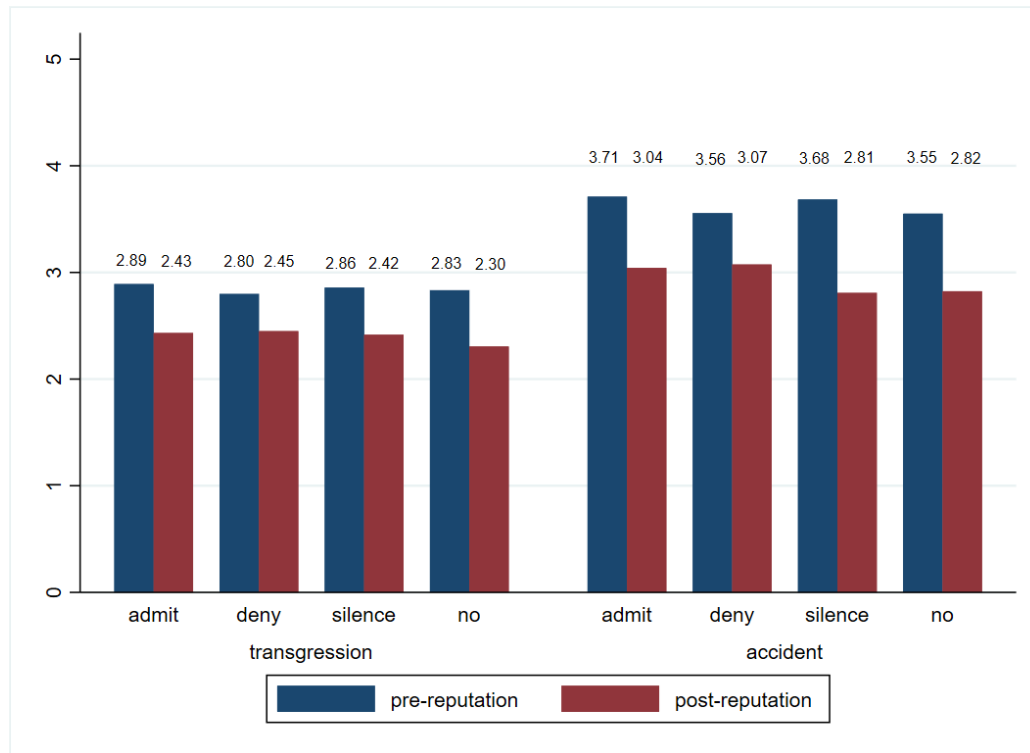


Figure 4.3: Mean comparison of pre and post reputation scores by strategic responses that are sub-grouped by failure types.

The results of the regression analyses are presented in Table 4. Using no action as a reference group, Model 1 showed that a denying response to the failure had a smaller effect on the respondents than on those who received no response. This result was statistically significant, but contrasted with Avoidance hypothesis (a), in that it showed that a denying response was more effective in the case of transgressions. Further, in Model 2, the results supported Avoidance hypothesis (b), that a denying response was significantly more effective in the case of an accident. These findings demonstrate that when agencies use different strategic responses, they can influence their reputations differentially, and in particular, a denying response was effective in the case of both types of failure, transgressions and accidents.

In addition, when we parse the types of failure and provide the results for each federal agency, Model 5 showed that the reputation judgments of respondents who received an admitting response and denying response decreased less than did those who received no response. However, even if both the VA and IRS were categorized as transgressions, Model 7 showed that none of the strategic responses differed significantly in the case of the IRS. With respect to the cases of the USPS and the CDC, where the agencies were categorized as accidents, Models 9 and 11 indicated that a denying response was a more effective way for these agencies to maintain their reputations in the eyes of citizens. The results showed that a denying response generally was the most effective way for agencies to deal with their failures.

Table 4.4: Regression analyses of the effect of response strategies on bureaucratic reputation

	Failure Types				Federal Agencies							
	Transgression		Accident		VA		IRS		USPS		CDC	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Admit	0.08 (0.07)	0.07 (0.07)	0.11 (0.08)	0.09 (0.08)	0.25** (0.10)	0.32*** (0.11)	-0.08 (0.09)	-0.10 (0.09)	0.12 (0.11)	0.12 (0.11)	0.11 (0.12)	0.06 (0.13)
Deny	0.16** (0.07)	0.24*** (0.07)	0.23*** (0.08)	0.21*** (0.08)	0.27*** (0.10)	0.52*** (0.11)	0.05 (0.09)	0.05 (0.09)	0.26*** (0.11)	0.24** (0.11)	0.20* (0.12)	0.18 (0.12)
Silence	0.09 (0.07)	0.16** (0.07)	-0.09 (0.08)	-0.08 (0.08)	0.05 (0.10)	0.25** (0.12)	0.11 (0.08)	0.11 (0.09)	-0.10 (0.10)	-0.08 (0.11)	-0.09 (0.11)	-0.08 (0.12)
No action												
Pre-reputation	0.63*** (0.02)	0.55*** (0.05)	0.59*** (0.03)	0.55*** (0.67)	0.58*** (0.04)	0.38*** (0.07)	0.66*** (0.03)	0.67*** (0.07)	0.65*** (0.05)	0.63*** (0.09)	0.52*** (0.05)	0.05*** (0.10)
Admit*Pre		-0.03 (0.07)		0.07 (0.10)		0.05 (0.10)		-0.06 (0.09)		0.00 (0.13)		0.16 (0.15)
Deny*Pre		0.17** (0.07)		0.08 (0.09)		0.38*** (0.09)		-0.02 (0.09)		0.07 (0.12)		0.08 (0.14)
Silence*Pre		0.15** (0.07)		-0.02 (0.09)		0.29*** (0.10)		0.01 (0.09)		-0.04 (0.13)		0.03 (0.14)
Constant	-0.23*** (0.05)	-0.26*** (0.05)	-0.18*** (0.06)	-0.17*** (0.06)	-0.39*** (0.07)	-0.53*** (0.08)	-0.08 (0.06)	-0.08 (0.06)	-0.18** (0.08)	-0.17** (0.08)	-0.19** (0.08)	-0.17* (0.08)
N	788	788	788	788	404	404	384	384	416	416	372	372
R ²	0.47	0.48	0.29	0.29	0.41	0.45	0.53	0.53	0.34	0.34	0.25	0.25

Note: p* < 0.1, p** < 0.05, p*** < 0.01, standard errors in the parenthesis

Bias Hypothesis

This study also tested the interaction effect between people's prior judgments and the denying response on their post-reputation judgments. In Table 4.4, Model 2 indicated that people who had a more positive reputation judgment were less likely to judge an organization's reputation more poorly when they received a denying response from agencies than when they received no response. These results supported the Bias hypothesis that people tend to believe information that confirms their prior beliefs. However, the results supported the hypothesis only in part because in Model 4, in the case of an accident, none of the strategic responses was associated significantly with people's prior reputation judgments. What we can see from these results is that people's prior judgments of the agencies influenced the effect of denying responses on their post-judgments only when the failure involved more organizational responses.

When we parse the types of failure according to agencies, Model 6 showed that both a denying response and strategic silence were more effective when people judged the VA's reputation more positively than when they received no response. However, in Models 8, 10, and 12, the findings indicated that the effect of people's prior reputation judgments was associated with none of the responses only in the case when the agencies were the IRS, USPS, and CDC. Thus, these results supported the Bias hypothesis only in part, in that people with more positive reputation judgments tended to believe the information that confirmed their prior beliefs only in the case of the VA.

6. Conclusions

The findings of this study demonstrated that the relevant use of organizations' strategic responses can mitigate the adverse effects of their failures on citizens' reputation judgments. The first reason for this is that citizens attribute the organizations' responsibility for the failure differently depending on their recognition of the cause, as attribution theory assumes (Weiner et al., 1980). However, while people had more positive pre-reputation judgments of the organizations that were assigned to the case of accidents than of transgressions, as we expected, the degree to which their post-judgments declined was greater in the case of accidents than transgressions, which contrasted with our expectation. These results indicated that people attribute more responsibility to organizations when the failure was committed by those of which they had higher expectations. These results are important, because they conflict with the assumption of the halo effect, that more positive reputations are a buffer against organizations' reputational damage (Coombs, 2006). The results actually supported a reverse halo effect, that people reacted more harshly to the wrongdoings of organizations with more positive reputations.

Secondly, this study's findings suggested that a particular type of strategic response is more effective for a particular type of failure. Specifically, a denying response was more effective in the case of both accidents and transgressions, which confirmed the symbolic model Coombs and Holladay (1996) suggested only in part. Further, this finding is in contrast to the assumption of communication research in the bureaucratic reputation literature, that organizations use an admitting response when there is a greater expectation of losing their reputations (Maor et al., 2012; Gilad et al., 2013). This can be explained based on two limitations of the research design: the first is that the federal agencies that

were selected in this research were exceptional and not representative, and the second was the manipulation of the experimental treatment. A possible argument for this finding is that the cases in this study were relatively lower profile and had not been main news headlines, as in the case of Hurricane Katrina, so that the severity of the case was not sufficiently strong to see the effect of an admitting response. For example, in the transgression cases, those of the VA and IRS, in which the agencies were blamed nationally for their responsibility, there is a possibility nonetheless that the cases might not be seen as ‘timely’ scandals with a high profile.

These limitations also affected the third finding of this study, the influence of people’s prior judgments on receiving strategic responses that affected their post-reputation judgments ultimately. The findings also partially support halo-effect of reputations on people’s receiving messages. What the findings of this study suggested is that people with more positive reputation judgments tend to believe agencies that deny their responsibility more only in the case of transgressions. This result may indicate that the halo effect only functions when an organization’s message confirms their beliefs in a more serious failure. However, there is another possible explanation, that the VA could be an exceptional case, given that its unique characteristics, such as recent scandals, mission, or respondents’ status as veterans, may have influenced these results.

Despite these limitations, the theoretical implications of this study are that it provides an empirical, interdisciplinary framework that links the communication, political science, and public administration literature to test the effect of presentational strategy on citizens’ reputation judgments using a survey experiment. These approaches extended the discussion of blame avoidance and the bureaucratic reputation literature by applying

Coombs and Holladay's (1996) symbolic approach that is grounded in attribution theory. The theoretical framework connects the two streams of research, blame avoidance and bureaucratic reputation, which has grown separately in the fields despite their close connection. Further, the research design opens the way to study the association between blame avoidance strategies and bureaucratic reputation empirically by employing an individual-level research design using a survey experiment. These efforts enrich the discussion of both blame avoidance and the bureaucratic reputation literature.

Further, this study's exploratory findings suggest to public managers that organizations' responses to an accusation are important in managing their relations with the public. The study's findings suggest that explaining is better than doing nothing about a failure in general, and emphasize the importance of providing explanations for an issue about which the public might wonder. The research suggests that public managers as well as public administration scholars focus more on the importance of presentation strategies that emphasize "communication" with citizens. Despite the importance of communication in citizen-state interactions, surprisingly few studies in the public administration literature have focused on the topic, which provides fertile ground for public managers to use.

An important topic for future research would be to test a similar work with a longitudinal design. While the design in this study can make only a cross-sectional comparison among agencies, it remains for future researchers to investigate the changes in the effectiveness of responses over a more extended period by conducting repeated experiments on the same panel of people. Further, future studies should include more federal agencies or different levels of government, which could overcome the limitations of this study, which included only four different federal agencies. Thirdly, the manipulation

of the failure types should be designed carefully to test Coombs and Holladay's (1996) symbolic approach. Research on different types of failure on the part of the same agencies could be one another manipulation that future researchers can implement.

One last suggestion this study would like to raise is that blame is discussed in the media in our everyday lives, and public agencies always respond to these issues in particular ways, which is important, but has been unexamined to date in the public administration literature. Managing external constituents is one of the major themes in public management, and the question of the way to do this should be given more attention. The presentational strategy is the first step that public agencies can use to restore their reputations and ultimately, maintain their legitimacy in the public's eyes when a failure occurs. This is a way for public agencies who are managing their resources and "...building and maintaining the level of public trust in government that makes democracy possible" (Goodsell, 2006, p. 633) to improve their selection and implementation of blame avoidance strategies.

CHAPTER 5:

DISCUSSION, LIMITATIONS, AND IMPLICATIONS

1. Introduction

The goal of this dissertation was to enhance our understanding of bureaucratic reputation in the eyes of citizens, particularly the way citizens formulate reputation judgments of U.S. federal agencies. The dissertation included three empirical studies on bureaucratic reputation and used the theoretical framework of accountability that views bureaucratic reputation as an external sanction on public organizations' behavior. Further, using the governance theory, three studies focused on the role of citizens, who are one of the key stakeholders and also the ultimate principals of public administration in the democratic governance system, in the relationships of public administration. Particularly from the perspective of NPM, citizens' reputation judgments were considered an element of public organizations' performance indicators that contains an overall and comprehensive evaluation of public organizations.

Further, three studies in this dissertation provided a relevant research design that could be applied to study bureaucratic reputation at the individual level. The first provided relevant tools to measure individuals' reputation judgments by developing a Bureaucratic Reputation Scale through standard scale development processes. The second explored individual characteristics that are associated with citizens' reputation judgments of federal agencies using national survey data. The third tested the relevance of reputation management skills, and focused specifically on communication strategies that affect citizens' reputation judgments. Survey experiments were used for empirical assessments. Those three empirical studies have important implications that enhance our understanding

of bureaucratic reputation by developing and implementing individual-level research designs that were absent in the current literature.

The following section summarizes the findings and includes the way the three empirical studies answered the dissertations' three major research questions that were presented in Chapter 1. This section recaps the findings of each study and states why those findings answered the research questions. Moreover, given the potential that the findings had weaknesses and limitations that may have threatened their validity, the third section describes the studies' methodological and theoretical limitations. Nevertheless, the studies in this dissertation are important and have the potential to contribute to the theoretical discussion of bureaucratic reputation, and further, to solve problems that practitioners face in the field of public administration. Thus, the fourth section discusses the theoretical implications of this dissertation, and the fifth provides practical implications. Building upon those implications, the final two sections offer recommendations for future studies.

2. Summary of Findings

The overall theme of the three studies in this dissertation was to study the formation and management of reputation by examining citizens' reputation judgments at the individual level. Based on this central theme, the dissertation developed the three major research questions stated in Chapter 1. The three studies conducted were designed to provide answers to those research questions; thus, this section reiterates the theoretical frameworks, research designs, and the way each study's results addressed those questions. Among the three, the first developed and validated the Bureaucratic Reputation Scale developed; the second provided empirical assessments of reputations' formation in the eyes

of citizens, and the third probed the management of reputations empirically by focusing on communication strategies. These studies are elaborated further in this subsection.

2.1 Measuring bureaucratic reputation: Scale development and validation

The purpose of the first study was to determine a way to measure bureaucratic reputation at the individual level. In response to the problem of the absence of relevant measurements, the first study's goal was to develop a scale that measures an individual's reputation judgment that can be used in survey studies. Scales refer to measurement instruments that are "...the collections of items combined into a composite score and intended to reveal levels of theoretical variables not readily observable by direct means" (DeVillis, 2003, p. 9). As bureaucratic reputation is composed of the collection of individuals' perceptions, developing and validating a scale is the appropriate way to capture this unobservable variable.

As its theoretical framework, this study adopted Carpenter's (2010) four dimensions of bureaucratic reputation, which include performance, morality, procedural fairness, and technical competence. By adding one more dimension, general reputation, the scale that this study developed included a total of five content domains of bureaucratic reputation. To develop and validate those standardized measurements of bureaucratic reputation, the study followed Devillis' (2003) standard scale development process. The scale, which was developed through the process of a literature review and expert review, also was tested empirically using a survey of U.S. citizens. The statistical reliability and validity tests confirmed that the 5- and 10-item scales that were developed in the first study were appropriate to use as survey tools to capture individuals' reputation judgments. Thus,

throughout these scale development processes, the study confirmed that the answer to the first research question, “*Can we measure bureaucratic reputation at the individual level?*” is indeed yes.

2.2 Predictors of citizens’ reputation judgments

The purpose of the second study was to examine the formulation of citizens’ reputation judgments. Bureaucratic reputation is formulated through interactions between public organizations and stakeholders. Focusing specifically on citizens as key stakeholders, the second study explored the way individual characteristics influence their reputation judgments. Understanding individual characteristics is important to manage bureaucratic reputation because they are stable predictors. Knowing those fundamental factors that determine citizens’ reputation judgments is a necessary step before developing different strategies to manage organizational reputation.

Thus, in the second study, individual characteristics, including demographics, socioeconomic status, geography, and ideologies, were assessed to determine their association with citizens’ reputation judgments of U.S. federal agencies. The theoretical framework of this study was grounded in policy preference frameworks that have been applied often to study public opinion (Page & Shapiro, 1990). The empirical assessment was based on data from the 2013 Pew Research Center that provided national survey data on 1,000 U.S. citizens. Among all of the variables, ideological factors, such as political ideology and trust in government, were associated strongly with a large number of federal agencies’ reputations. Thus, the answer to the second research question, “*What factors are associated with citizens’ reputation judgment?*” is that individual characteristics,

particularly ideological factors, including political ideology and trust in government, shape citizens' reputation judgments.

2.3 Reputation management and communication strategies

The third study's purpose was to test empirically the effect of communication strategies on citizens' reputation judgments. Bureaucratic reputation is not static, but often unstable and variable (Maor 2015), indicating that relevant management can improve public organizations' reputations and protect them from reputational threats. Of the two major approaches to study reputation management, selective responses and communication strategies, the third study focused specifically on agencies' reputation management using communication strategies, and assessed empirically the effect of communication strategies on citizens' reputation judgments after policy and administrative failures.

Using blame avoidance as a theoretical framework and developing models based on the attribution and cognitive dissonance theories, the third study attempted to determine whether 1) using different strategies can influence citizens' reputation judgments after agencies' failure, and 2) individuals' pre-reputation judgments interact with those strategies' effects on their post-reputation judgments. The empirical assessment was based on survey experiments that were conducted on a sample of approximately 1,000 U.S. citizens. The results of the study suggested that problem denial was more effective in protecting U.S. federal agencies' reputation than are other strategies, including problem admission, nonresponse, or no action at all. Further, depending on the types of failure, people's pre-judgments interacted with the strategies' effects. Thus, the answer to the third

research question, “*Can public organizations manage citizens reputation judgments?*”, particularly through communication strategies, is “yes.”

3. Limitations

3.1 Methodological limitations

The three studies in this dissertation included methodological limitations. Firstly, the citizen surveys relied on an online sample of citizens that may not be fully representative of the general population. In particular, the first study relied on voluntary, non-probability samples from CivicPanel, which have the potential to limit the ability to generalize the results. Further, such a sampling strategy may cause sampling bias, particularly coverage bias that is related to the possibility of differences between the sample and population overall (Remler & Van Ryzin, 2010). Thus, it was necessary to test the robustness and reliability of the Bureaucratic Reputation Scale by using different samples of citizens. While the citizen sample in the third study was more representative, as the Qualtrics Panel was asked to use a quota that matched the proportion of the national population, the results should be interpreted with caution nonetheless because of the possible bias associated with the sample.

An additional limitation of the individual-level studies in this dissertation is that it is possible that the scale and the assessments may not work well in cases in which citizens have limited knowledge about, or experiences with, a government agency. The study in Chapter 2 tested the Bureaucratic Reputation Scale with three fairly well known U.S. federal agencies (NASA, FDA, and IRS) but, of course, the scale may not be as reliable or valid if used to assess the reputations of agencies about which the audience has more

limited understanding or awareness. Further, in Chapter 3, the low R-squares (around 0.1) might occur due to citizens' non-attitudes toward federal agencies that may be caused by limited knowledge or awareness. These findings suggest further research of predictors that may be more strongly associated with citizens' judgments.

The additional limitation that should be noted is that the findings from the three studies were based on only a small number of U.S. federal agencies, which makes it difficult to generalize the results to other U.S. federal agencies, other levels of government, and other countries and national contexts. Bureaucratic reputation is a concept that is not confined to federal agencies, as previous studies have focused largely on the FDA, for example; thus, it has great potential to be applied to other levels of government. For example, the state and local government levels, where the relationship between the government and citizens is close, might have different predictors that affect citizens. Further, international governments with diverse cultural and social backgrounds could also lead to results that differ from the findings in these three studies. The only way to apply the research design of this dissertation to other levels of governments is to test the findings' external validity and offer practitioners more ideas about the way particular groups of people perceive their organizations.

The final methodological limitation of all three studies in this dissertation is that they allowed only a cross-sectional comparison among agencies. While this approach allows us to capture the difference in reputations among agencies in the eyes of citizens, it cannot explain the changes in agencies' reputation over time. With a longer time-span, we could see the changes in citizens' reputation judgments, and the dynamics that depend on external changes, such as scandals or failures. This longitudinal approach is important,

particularly in bureaucratic reputation research, because the definition of reputation implies audiences' perceptions that are built historically among a network of multiple audiences (Carpenter, 2010). Many longitudinal studies in the current bureaucratic reputation literature have used media coverage; however, more longitudinal studies should be conducted on bureaucratic reputation that use survey research to enhance our understanding of the historical aspects of reputation, which are characteristics essential to understand the mechanisms fully.

3.2 Theoretical limitations

It is important to point out the theoretical limitations of this dissertation as well. First, while individual-level measures of bureaucratic reputation should be distinguished from other similar citizen perception concepts, such as citizen satisfaction or trust in government, this dissertation does not provide empirical evidence of the distinction. As stated in Chapter 1, following Van de Walla (2010), identifying citizen satisfaction is a measure too narrow, as it focuses on public organizations' service delivery alone, while trust in government is a measure too broad to measure at the institutional level. Theoretically, citizens' reputation judgments could be citizens' perceptions of organizations. However, despite the theoretical discussions of that differentiation, this dissertation does not provide an empirical assessment of the distinction of those similar concepts. More tests should be conducted on the differences among those similar measures to strengthen the validity of the Bureaucratic Reputation Scale, and further, to establish the reputation literature more firmly in the public administration literature.

Another important limitation is that the samples used to develop the Bureaucratic Reputation Scale and conduct the empirical tests in the dissertation were citizens when, importantly, the theory of bureaucratic reputation references a wider variety of audiences, including political and judicial authorities, interest groups and civic associations, professional and scientific expertise, the media, and the mass public (Carpenter, 2010). Tests that use only citizens can lead to the problem of external validity, as Carpenter's (2010) definition of bureaucratic reputation implies comprehensive perceptions of a network of multiple audiences. This problem could be overcome by using the Bureaucratic Reputation Scale with other populations (audiences) because its wording was designed to be general so that it could be used in that way.

Further, three studies in this dissertation relied on unidimensional reputation measures, despite the fact that Carpenter's (2010) definition relied on multi-dimensional measures. For example, the first and third studies used the Bureaucratic Reputation Scale, which uses a composite score of five items to represent the concept, while the second study used favorability as a proxy of organizational reputation. As the bureaucratic reputation literature is grounded in the multidimensional nature of reputations, there are possible limitations in using a one-dimensional measure. More discussion of the multidimensionality of bureaucratic reputation can be found in Overman, Buisuioc, and Woodman's (2020) work. Nevertheless, this study's justification of a one-dimensional measure is that the Bureaucratic Reputation Scale represents five content domains, performance, morality, procedure fairness, technical competence, and general, which is built upon Carpenter's (2010) work.

Lastly, three studies in this dissertation do not fully explain or control factors that may affect citizens' reputation judgments. One of the reasons for this was that the purpose of the first study was to develop a scale, the second was to explore the association between individual characteristics and citizens' reputation judgments, and the third was to test the hypothesis using an experimental setting that eliminates the concern about causality (Remler & Van Ryzin, 2010). Despite the fact that those studies did not require other covariates that might affect the relationships among the variables to be controlled, still, the results in this study should be interpreted with cautions. For example, the results in Chapter 2 and Chapter 3 regarding the association between political ideology and citizens' reputation judgments are contradictory. Specifically, while Chapter 2 suggested that political ideology and reputation judgments were unrelated, Chapter 3 supports such relationship. This can be explained in part perhaps by the fact that agencies seem to have different appeals depending on political ideology. The reputations of the FDA, IRS, and NASA, which were tested in Chapter 2, were insignificantly related to respondents' political ideology (except IRS) in the results of Chapter 3. The reputations of the EPA, DOE, and DOD were significantly related to people's political ideology in Chapter 3. One of the possibilities is that the FDA, NASA, and IRS are less politically sensitive to political ideology in citizens' judgments. Another possibility is that although the purpose of Chapter 3 was to explore the way individual characteristics shape citizens' reputation, the results may differ depending on the inclusion of covariates.

4. Implications for Future Agenda

As bureaucratic reputation has emerged as a fundamental concept in public administration because of its implications for bureaucratic power, resources, and politics, the need to measure reputations and understand their formation and management has increased. Three studies in this dissertation developed and validated the new bureaucratic reputation measurement, explored predictors, and assessed the strategies used to manage reputation empirically. Although they have various methodological and theoretical limitations, the three studies provide research frameworks to study bureaucratic reputation at the individual level empirically. Further, practically, the results revealed the abstract concept of reputations more clearly so that practitioners have the opportunity to measure and learn the way to manage public organizations' reputations more clearly. More specific implications of each study are discussed below.

4.1. Implications for theory and research

The three empirical studies in this dissertation have some implications for bureaucratic reputation researchers. First, their findings provide a way to study bureaucratic reputation at the individual level. Most of the research in the current bureaucratic reputation literature has studied reputations at the organizational level, and little attention has been paid to studying the concept at the individual level. As described in Chapter 1, this study followed the approach of methodological individualism—that social phenomena can be understood fully only when the holistic level approach is complemented by the individual level approach, which provides tools and research frameworks that can be applied to study reputations at that level.

For example, the Bureaucratic Reputation Scale will be useful for researchers who are interested in conducting surveys or survey experiments and require a standardized measure of bureaucratic reputation. Importantly, the scale has wording sufficiently general and flexible to be used in studies of various agencies, including not only other U.S. federal agencies, but state and local government agencies as well as those in other countries. The use of a standardized measurement like the Bureaucratic Reputation Scale in surveys and other studies of bureaucratic reputation can facilitate the comparison of results, establishment of empirical consistencies, and in turn, development of theory across agencies, countries, and policy contexts.

Further, the individual-level approach of the research designs in this dissertation also has theoretical implications that shed light on an audience-based approach that Maor (2016) emphasized as a “missing area in the bureaucratic reputation framework.” According to the author, audiences refer to observers of agencies, and when they observe, many factors, including “...prior knowledge, goals, mental frames, heuristics, distraction, motivation, emotion, and others,” may affect them (p. 82). While research has focused predominantly on the agency-side and the factors that affect agencies’ behavior, the audience-side is also important, as bureaucratic reputation is an interaction between agencies and audiences. As the audiences’ perceptions have characteristics that are unobservable, diverse, and variable, they can be understood from both “social” and “psychological” perspectives (Maor, 2016). This indicates that those perceptions can be assessed in empirical studies with individual-level approaches. The tools and research designs that this dissertation provides could be a relevant foundation upon which future studies can build to study bureaucratic reputation with the audience-based approach.

Particularly, Chapter 3 provide the way individual characteristics shapes citizens' reputation judgments of federal agencies. The findings that citizens do not have identical views of federal agencies has theoretical implications; the public's views of agencies vary depending on region, ideology and socioeconomic characteristics, as well as race and gender. While R^2 was relatively low, still the findings provide an important foundation to the extent that future studies could expect individual characteristics to be stable predictors and less likely to be changeable according to external shocks. By understanding these fundamental characteristics of citizens' reputations, studies on bureaucratic reputation can be further developed.

Secondly, three studies in this dissertation relied on cross-sectional research designs, which was described as a limitation in the previous section, as it provides a shortage of answers relevant to the longitudinal and evolutionary natures of reputations. Nevertheless, the findings offer certain insights into the distinctive reputations across U.S. federal agencies in the eyes of citizens and factors associated with reputation judgments. The results showed that some agencies are judged to have better reputations than others, and this variation is associated with citizens' demographic, socioeconomic, geographic, and ideological characteristics. For example, in Chapter 3, the CDC was rated as having the best reputation among the federal agencies, and its reputation was associated positively with income and also region (Northeast and South). The IRS was rated as having the worst reputation among the 12 agencies in this study, although the results showed that liberals tend to judge the IRS' reputation more positively than do conservatives. These results suggest that there is variation in citizens' reputation judgments of different federal agencies,

and such variations provide empirical grounds to test the factors that affect those variations in future research.

Further, the research design in Chapter 4 provides future researchers with an idea that could be applied to a longitudinal setting. From the findings, we saw that there was a significant effect of communication strategies on citizens' reputation judgments depending on the types of failure that public agencies experienced. In the longitudinal setting, this research design can be tested repeatedly to see the way people's responses change over time passes. For example, based on the research design in Chapter 4, future studies can recruit a panel of people and give the first treatment (including public agency scenarios and responses) at time 1, the second at time 2, and the third at time 3. With this design, we can obtain results that are more realistic and similar to real-world settings, where people are exposed repeatedly in media coverage to similar news articles about agencies' failures—and determine the way people react differently to the public organizations' responses as time passes.

Thirdly, the theoretical framework of this dissertation applied the intersection of multiple theories in various fields that could benefit discussions in the public administration literature. For example, in Chapter 3, the research applied the policy preference framework that has been used largely in the public opinion literature from the perspective of political science. By incorporating that framework into the bureaucratic reputation framework and identifying the relevant results from the frameworks' empirical assessment, this study provided a discussion of citizens' reputation judgments from the perspective of public opinion. This effort was nascent in the public administration literature

and also has great potential for future researchers to adopt similar frameworks to study the way individual characteristics shape citizens' reputation judgments.

Further, the theoretical framework of Chapter 4 suggested a way to incorporate the literature on public administration, communication, and political science. The broad setting of the theoretical framework was based on Blame Avoidance, which is a popular theory in both political science and public administration. Particularly, the blame avoidance theory's focus on presentation strategies has been tested rarely at the individual level in the public administration literature because of the lack of relevant hypothetical models. By applying the model Coombs and Holladay (1990) developed in the communication literature, this study provided relevant multidisciplinary testable models that also could be used in future studies in the field of public administration. The implication of this research is that it provides multidisciplinary models to conduct empirical tests of the important theoretical discussions in the public administration literature.

The fourth theoretical implication is that these studies, particularly the one in Chapter 2, provide opportunities for bureaucratic reputation researchers to discuss unidimensional vs. multidimensional reputation. The current literature mostly relies on multidimensional approach, viewing that bureaucratic reputation has four multiple dimensions—performance, morality, procedural fairness, and technical competence. However, in business literature, still there are ongoing debates about whether reputations are unidimensional or multidimensional (Walsh et al., 2007), which suggests that the bureaucratic reputation literature also needs more discussions of dimensionality. The discussion about reputations' dimensionality has theoretical importance, as it calls for discussion of the way to use the reputation measure.

Although the Bureaucratic Reputation Scale developed in Chapter 2 is unidimensional, it does not deny the existence of multiple “dimensions” of reputation. Rather, the scale adopts the view that reputation tends to be a single concept with five different content “domains.” The reasoning behind this approach is rooted in the fundamental discussion over unidimensional vs. multidimensional measures. Often, unidimensional and multidimensional measurements for a single concept coexist; the difference lies in the fact that a multidimensional measurement places more emphasis on different origins of the concepts, while a unidimensional measurement assumes that respondents cognitively aggregate their judgments of each content domain into a single integrated response (Ironson, Smith, Brannick, Gibson, Paul, 1989; Wright, Christensen, Pandey, 2013; Van Engen, 2018). A multidimensional scale has limitations as a useful and applicable measurement, moreover, including fatigue and response bias attributable to its considerable length. Thus, this study suggests that a unidimensional measure (Bureaucratic Reputation Scale) has advantages both conceptually and empirically. Nevertheless, future researchers are expected to have a continuing discussion of the dimensionality of bureaucratic reputation.

Lastly, as a closing remark, I would like to suggest that future researchers study bureaucratic reputation in a broader theoretical setting. For example, few studies have investigated reputation from the Resource-based View or its role in networks of service provision where diverse sector actors participate in delivering public services. Reputations are resources essential for public organizations to achieve greater autonomy, power, and legitimacy; at the same time, reputation is a sanction that guides public organizations to engage in more accountable behavior. Thus, reputations are relevant lenses that could

enhance our understanding of public organizations' behavior in today's complex society and provide essential knowledge that leads to more democratic governance systems.

Further, there are possible predictors in addition to those (individual characteristics and communication strategies) that are likely to influence citizens' reputation judgments. In addition to the importance of political knowledge of citizens that might affect their reputation judgments of federal agencies, the leadership of an agency may be an relevant predictor, as people's knowledge of agencies may be related closely to their familiarity with who heads an agency (Maor, 2016). For example, the scandal involving Scott Pruitt, an EPA administrator, may have affected the EPA's actual reputation. Identifying the relations between an agency's head and an agency's reputation has great potential for the future research agenda.

4.2. Implications for policy and practice

The findings from this dissertation have certain policy and practical implications—the way public organizations or managers can use knowledge. This section begins with three primary challenges Carpenter and Krause (2012) raised for public administrators that are fundamental to governance, and the way this dissertation could provide certain ideas to address two among the three: 1) the way to maintain broad-based support for an agency and its activities; 2) the way to steer a vessel amid hazardous shoals (enemies and potentially disaffected supporters), and 3) the way to project a judicious combination of consistency and flexibility.

4.2.1 The way to maintain broad-based support for an agency and its activities

The first challenge is for public agencies to manage audiences' support for them and their activities, which is fundamental environmental input that influences agencies' policy outputs according to the general systems theory (Rainey, 2009). In the reality of democratic settings, the operation of public administration is not isolated, but is engaged actively in interactions with the political environment—multiple audiences, including elected officials, clientele groups, the media, policy experts, and ordinary citizens (Carpenter & Krause, 2012). However, the support from the audiences is unobservable and ambiguous in this setting, and thus, public organizations need a measurement that reveals their audiences' support.

This study provides relevant tools (Chapter 2) that public organizations and managers can use to capture their audiences' support. Particularly, the tools are designed to target individual-level audiences. The Bureaucratic Reputation Scale offers public managers a great advantage to investigate multiple audiences' diverse perceptions of public organizations, where the public managers must be aware of their audiences, which have different demands and expectations (Carpenter & Krause, 2012). Relying on individual-level measures is the only way they can determine the differences among diverse groups of audiences. Media coverage is often considered a preferred method to understand audiences' support of organizations. This method has strength in revealing the comprehensive perception of multiple audiences, but it has a weakness in revealing different perspectives among different groups. Thus, the way suggested to solve this problem is for public managers to use new individual-level tools to gauge multiple

audiences' diverse perceptions and identify strategically which audiences they should prioritize.

4.2.2 The way to steer a vessel amid hazardous shoals (enemies and potentially disaffected supporters)

4.2.2.1 Conflict of audiences' interest

Another challenge that public administration confronts in a democratic system is that audiences experience conflicts of interest, and public agencies are responsible for all of those audiences. One major difference between the public and private sectors is that the goal of public organizations is not just profit, as in private organizations, but the pursuit of more diverse values, including fairness, morality, and competencies. This implies that public organizations should be aware of the conflicts among audiences, and further, the component groups of those audiences that are complex and difficult to resolve. Failing to cope with those conflicts may lead governments to follow obscure paths and take a longer time to accomplish the work necessary to solve social problems in reality.

The first thing that public managers should do is determine the factors that affect those audiences' differences in preference or evaluation. While the studies in this dissertation focused only on ordinary citizens among many other audiences, the findings in Chapter 3 suggested that public managers have different groups within one type of audience that can hold various perceptions of organizations. For example, the results showed that people who are more educated are more likely to favor NASA, but are less likely to prefer HHS and DOE. This finding may suggest that the way HHS and DOE operate their policies may be supported by less educated people, but not by people who are more educated. This can be explained by the direction of the Obama Administration, which

targeted less privileged people with better healthcare and education opportunities. Further, the findings also suggested that liberals have contrasting attitudes toward “liberal agencies” and “conservative agencies,” as described in Chapter 2. In reiterating the implications for future research, it would be interesting to see the change in the results under President Trump and a Republican-controlled Congress. As the empirical findings actually showed the diverse views among different subgroups of citizens, understanding those conflicts and contrasting views that depend on audiences’ characteristics is the first step public managers must take to prepare for conflict resolution.

4.2.2.2 Coping with failures

Public agencies experience failures in their operation inevitably, which lead to opposition to the agencies and ultimately decrease audiences’ support for government. As described in Chapter 4, failure is based on audiences’ perceptions, and careful responses can protect agencies from losing their reputation (van de Walle, 2010; Moynihan, 2012). This indicates that public managers need certain systematic empirical grounds with which they can calibrate their strategic responses to audiences to protect their organizations’ reputations.

Chapter 4 opens the discussion for the use of communication strategies in public administration. Such strategies emphasize using an argument to cope with failure that has an immediate effect on audiences’ perceptions of that failure (Hood, 2012). It offers ideas to practitioners, in that the way they present themselves can influence audiences’ perceptions, and ultimately, their reputations. Of course, the real-world setting is more complex, and simply admitting or denying the failure is insufficient; thus, additional

research should be performed to reveal more of the complexity in the responses. For example, as Hood, Jennings, Dixon, Hogwood, and Beeston (2009) found in his study, the effects of the response vary depending on the combinations of those responses in a timely order (e.g., problem denial + admission). Further, the effect of strategic responses can differ in a setting where the failure types are not simply local, but national level scandals. Those will provide more information to practitioners, and the findings in Chapter 4 have a practical implication, in that strategic responses are important to manage relationships with audiences, particularly when the situation is negative.

5. Closing Remarks

Over the past several decades, we have witnessed changes in the governance system of public administration—they involve more actors in service delivery, citizens are involved more as evaluators of the services, and government-audience interactions are complex because of the conflicting demands and needs of audiences in a democratic society. Indeed, the changes in the systems put the traditional public agencies' accountability into question and require more flexible and well-functioning sanctions and incentives that fit contemporary society. This dissertation suggested that bureaucratic reputation, which is composed of diverse perceptions of organizations, could be an accountability mechanism that involves a network of multiple audiences and allows a flexible response to different demands from society. These governance discussions rest on accountability mechanisms and NPM, which also brings citizens' perceptions into the center of reputation formation and management.

Bureaucratic reputation is an essential concept that should be discussed in public administration. What is expected in this dissertation is that the theoretical and methodological efforts in the studies have generated a greater focus on the importance of reputations at the individual level and increased attention to reputations' formation and management. As Carpenter and Krause (2012) stated that the way "...organizational reputations are formed and subsequently cultivated is fundamental to understanding the role of public administration in a democracy" (p. 26).

REFERENCES

- Abramowitz, A. I. (1975). Name familiarity, reputation, and the incumbency effect in a congressional election, *The Western Political Quarterly*, 28(4), 668–684.
- Acock, A. C. (2012). What to do about missing values. In C. Harris (ed.), *APA handbook of research methods in psychology, Vol 3: Data analysis and research publication*, Washington, DC: American Psychological Association, 27–50
- Alford, J. R. (2001). We're all in this together: The decline of trust in government, 1958-1996. In J. Hibbing and E. Theiss-Morse (eds.) *What is it about government that Americans dislike?* Cambridge: Cambridge University Press, 28-46.
- Alexander, J. (1987). *The Micro-Macro Link*. Berkeley, CA: University of California Press.
- Allard S. W. & Cigna, J. (2008). Access to social services in rural America: the geography of the safety net in the rural west. RUPRI Rural Poverty Research Center Working Paper: 08-01.
- Allison, G. (1982). Public and private management: Are they fundamentally alike in all unimportant respects? In F. S. Lane (ed.), *Current Issues in Public Administration*. New York: St. Martin's Press.
- Althoff, P. & Greig, W. H. (1977). Environmental pollution control: Two views from the general population. *Environment and Behavior*, 9(3), 441-456.
- Arrow, K. J. (1994). Methodological individualism and social knowledge. *The American Economic Review*, 84(2), 1-9.
- Baekgaard, M., Christensen, J., Dahlmann, C., Mathiasen, A., & Petersen, N. (2019). The role of evidence in politics: Motivated reasoning and persuasion among politicians. *British Journal of Political Science*, 49(3), 1117-1140.
- Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99-120.
- Bennett, R. & Gabriel, H. (2003). Image and reputational characteristics of UK charitable organizations: An empirical study. *Corporate Reputation Review*, 6 (3), 276-89.
- Bouckaert, G., & van de Walle, S. (2003). Comparing measures of citizen trust and user satisfaction as indicators of “Good Governance”: Difficulties in linking trust and satisfaction indicators. *International Review of Administrative Sciences*, 69(3), 329–343. <https://doi.org/10.1177/0020852303693003>
- Bovens, M. (2007). Analyzing and assessing accountability: A conceptual framework. *European Law Journal*, 13(4), 447-468.

- Boyd, B. K., Bergh, D. D., & Ketchen Jr. D. J. (2010). Reconsidering the reputation—performance relationship: A resource-based view. *Journal of Management*, 36 (3), 588-609.
- Bromley, D. B. (2000). Psychological aspects of corporate identity, image, and reputation. *Corporate Reputation Review*, 3(3), 240–252.
- Brooks, M. E., Highhouse, S., Russell, S. S., & Mohr, D. C. (2003). Familiarity, ambivalence, and firm reputation: Is corporate fame a double-edged sword? *Journal of Applied Psychology*, 88(5), 904–914.
- Brummett, B. (1980). Towards a theory of silence as a political strategy. *The Quarterly Journal of Speech*, 66, 289-303.
- Brummett, B. (1980). Towards a theory of silence as a political strategy. *The Quarterly Journal of Speech*, 66, 289–303.
- Bryson, J. M., Ackermann, F., & Eden, C. (2007). Putting the resource-based view of strategy and distinctive competencies to work in public organizations. *Public Administration Review*, 67(4), 702-17.
- Bussemeyer, M. R., Goerres, A. and Weschle, S. (2009). Attitudes towards redistributive spending in an era of demographic ageing: The rival pressures from age and income in 14 OECD countries. *Journal of European Social Policy*, 19(3), 195-212.
- Busuioc, E. M. (2016). Friend or Foe? Inter-agency cooperation, organizational reputation, and turf. *Public Administration*, 94 (1), 40–56.
- Busuioc, E. M., & Lodge, M. (2016). The Reputational basis of public accountability. *Governance*, 29(2), 247–263.
- Byrkjeflot, H. (2015). Driving forces, critiques, and paradoxes of reputation management in public organizations. In M. Maor, & A. Waeraas, *Organizational reputation in the public sector* (pp. 54-73). New York, NY: Routledge.
- Campbell, D. T., and Fiske, D W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin*, 56(2), 81.
- Carmines, E. G., and Zeller, R A. (1979). *Reliability and validity assessment*. Vol. 17. Thousand Oaks, CA: Sage publications.
- Carpenter D. P. (2010). *Reputation and power: Organizational Image and Pharmaceutical Regulation at the FDA*. Princeton: Princeton University Press.
- Carpenter, D. P. (2001). *The forging of bureaucratic autonomy: Reputations, networks, and policy innovation in executive agencies, 1862-1928*. Princeton, NJ: Princeton University Press.

- Carpenter, D. P. (2002). Groups, the media, agency waiting costs, and FDA drug approval. *American Journal of Political Science*, 46(3), 490–505.
- Carpenter, D. P., & Krause, G. A. (2012). Reputation and public administration. *Public Administration Review*, 72(1), 26–32.
- Caruana, A. (1997). Corporate reputation: concept and measurement. *Journal of Product & Brand Management*, 6(2), 109-118.
- Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods & Research*, 36(4), 462-494.
- Chong, D., & Druckman, J. N. (2007). Framing theory. *Annu. Rev. Polit. Sci.*, 10, 103–26.
- Christensen, T., & Lægreid, P. (2005). Trust in government: The relative importance of service satisfaction, political factors, and demography. *Public Performance & Management Review*, 28(4), 487-511.
- Christensen, T., & Lodge, M. (2018). Reputation management in Societal Security: A comparative study. *American Review of Public Administration*, 48(2), 119-132.
- Chun, R. (2005). Corporate reputation: Meaning and measurement. *International Journal of Management Reviews*, 7(2), 91–109.
- Cook, F. L. (1979). *Who should be Helped? Public Support for Social Services*. Thousand Oaks: Sage Publications.
- Coombs, T. (1995). Choosing the right words: The development of guidelines for the selection of the “appropriate” crisis-response strategies. *Management Communication Quarterly*, 8(4), 447–476.
- Coombs, T. W., & Holladay, S. (2006). Unpacking the halo effect: reputation and crisis management. *Journal of Communication*, 10(2), 123-137.
- Coombs, W. T., & Holladay, S. (1996). Communication and attributions in a crisis: An experimental study in crisis communication. *Journal of Public Relations Research*, 8(4), 279-295.
- Daft, R. (1983). *Organization theory and design*. New York, NY: West .
- Darley, J., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of personality and Social Psychology*, 44(1), 20-33.
- Darley, J., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of personality and Social Psychology*, 44(1), 20-33.
- Davies, G., Chun, R., Silva R. V. D., & Roper, S. (2001). The personification metaphor as a measurement approach for corporate reputation. *Corporate Reputation Review*, 4(2), 113-27.

- Dawar, N., & Pillutla, M. (2000). Impact of product-harm crisis on brand equity: The moderating role of consumer expectations. *Journal of Marketing Research*, 37(2), 215-226.
- Dawar, N., & Pillutla, M. (2000). Impact of product-harm crisis on brand equity: The moderating role of consumer expectations. *Journal of Marketing Research*, 37(2), 215-226.
- DeVellis, R. F. (2003). *Scale development: Theory and applications*. (2nd edition). Vol. 26. Thousand Oaks, CA: Sage publications.
- Dowling, G. and Gardberg, N. (2012). Keeping Score: The Challenges of Measuring Corporate Reputation. In M. L. Barnett & T. G. Pollock (Eds.), *The Oxford Handbook of Corporate Reputation*. Oxford: Oxford University Press.
- Dowling, G. R. (2001). *Creating corporate reputations*. Oxford, UK: Oxford University Press.
- Eismeyer, T. J. (1982). Public preferences about government spending: Partisan, social, and attitudinal sources of policy differences. *Political Behavior*, 4(2), 133–145.
- Feldman S (1988) Structure and consistency in public opinion: The role of core beliefs and values. *American Journal of Political Science*, 32(2): 416-440.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
- Finer, H. (1941). Administrative responsibility in democratic government. *Public Administration Review*, 1, 335-350.
- Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. (2000). The Affect Heuristic in Judgments of Risks and Benefits. *Journal of Behavioral Decision Making*, 13(1), 1–17.
- Fishbein, M. leek Ajzen (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Boston, MA: Addison-Wesley Publishing Company
- Fombrun, C. and Shanley, M. (1990). What’s in a name? Reputation building and corporate strategy. *The Academy of Management Journal*, 33 (2), 233–58.
- Fombrun, C. J. (2012). The building blocks of corporate reputation: Definitions, antecedents, consequences In M. L. Barnett & T. G. Pollock (Eds.), *The Oxford handbook of corporate reputation*: 94–113. Oxford: Oxford University Press.
- Fombrun, C. J., & Van Riel, C. B. M. (2003). *Fame & fortune: How successful companies build winning reputations*. Upper Saddle River, NJ: Prentice Hall.
- Fombrun, C. J., Gardberg, N. A., & Sever, J. M. (2000). The Reputation QuotientSM: A multi-stakeholder measure of corporate reputation. *Journal of Brand Management*, 7(4), 241–255. <https://doi.org/10.1057/bm.2000.10>

- Fombrun, C.J. (1996), *Reputation: Realising value from the corporate image*. Cambridge, MA: Harvard Business School Press,.
- Fraile, M. & Ferrer, M. (2005). Explaining the determinants of public support for cuts in unemployment benefits spending across OECD countries. *International Sociology*, 20(4), 459-481.
- Frederickson, G. H. (2205). What happened to public administration? Governance, Governance everywhere. In E. Ferlie, L. Lynn Jr, L. Lynn, & C. Pollitt, *The Oxford Handbook of Public Management* (pp. 282-304). Oxford, UK: Oxford University Press.
- Friedrich, C. J. (1940). Public policy and the nature of administrative responsibility. In C. J. Friedrich, & E. Mason, *Public policy* (pp. 3-23). Cambridge, MA: Harvard University Press.
- Friedrich, C. J. 1940. Public policy and the nature of administrative responsibility. *Public*, 3-24.
- Gilad, S., Maor, M., & Bloom, P. B.-N. (2013). Organizational reputation, the content of public allegations, and regulatory communication. *Journal of Public Administration Research and Theory*, 25(2), 451-478.
- Goodsell, C. T. (2006). A new vision for public administration. *Public Administration Review*, 66(4), 623-635.
- Goodsell, C. T. 2014. *The new case for bureaucracy*. Washington DC: CQ Press.
- Griffin, E. (1994). *A first look at communication theory* (2nd ed.). New York, NY: McGraw-Hill.
- Grimmelikhuijsen, S., & Knies, E. (2017). Validating a scale for citizen trust in government organizations. *International Review of Administrative Sciences*, 83(3), 583-601
- Grimmelikhuijsen, S., Jilke, S. Olsen, A. L. & Tummers, L. (2016). Behavioral public administration: Combining insights from public administration and psychology. *Public Administration Review*, 77 (1), 45-56
- Grimmelikhuijsen, S., Jilke, S., Olsen, A. L., & Tummers, L. (2016). Behavioral public administration: Combining insights from public administration and psychology. *Public Administration Review*, 77(1), 45-56.
- Hall, R. (1992). The strategic analysis of intangible resources. *Strategic Management Journal*, 13, 135-144.
- Hart, D. K. (1984). The virtuous citizen, the honorable bureaucrat, and" public" administration. *Public Administration Review*, 44, 111-120.
- Hayek, F. v. (1942). Scientism and the Study of Society I. *Economica*, 9, 267–91.

- Hetherington, M. J. 2005. *Why trust matters: Declining political trust and the demise of American liberalism*. Princeton: Princeton University Press.
- Hood, C. (2012). The blame game: Spin, bureaucracy, and self-preservation in government. *West European Politics*, 35.
- Hood, C. C. (2001). *The blame game: Spin, bureaucracy, and self-preservation in government*. Princeton, NJ: Princeton University Press.
- Hood, C. C. (2003). The risk game and the blame game. *Government and Opposition*, 37(1), 15-37.
- Hood, C., Jennings, W., Dixon, R., Hogwood, B., & Beeston, C. (2009). Testing times: Exploring staged responses and the impact of blame management strategies in two examination fiasco cases. *European Journal of Political Research*, 48(6), 695-722.
- Ingold K and Leifeld P (2014) Structural and institutional determinants of influence reputation: A comparison of collaborative and adversarial policy networks in decision making and implementation. *Journal of Public Administration Research and Theory*, 26(1), 1-18.
- Ironson, G. H., Smith, P. C., Brannick, M. T., Gibson, W. M., & Paul, K. B. (1989). Construction of a job in general scale: A comparison of global, composite, and specific measures. *Journal of Applied Psychology*, 74(2), 193-200.
- Itami, H. (1987). *Mobilizing invisible assets*. Cambridge, MA: Harvard University Press.
- James, O., & Van Ryzin, G. (2017). Motivated reasoning about public performance: An experimental study of how citizens judge the Affordable Care Act. *Journal of Public Administration Research and Theory*, 27(1), 197-209.
- James, O., Jilke, S., Petersen, C., & Van de Walle, S. (2015). Citizens' blame of politicians for public service failure: Experimental evidence about blame reduction through delegation and contracting. *Public Administration Review*, 76(1), 83-94.
- Jilke, S. (2016). Microfoundations, microfoundations: Addressing 'Big Questions' through a behavioral approach toward public administration. Concept paper: Minnowbrook 50th Anniversary Conference.
- Jilke, S., Van Ryzin, G. G., & Van de Walle, S. (2015). Responses to decline in marketized public services: An experimental evaluation of choice overload. *Journal of Public Administration Research and Theory*, 26(3), 421-432.
- Jilke, S., Van Ryzin, G., & Van de Walle, S. (2015). Responses to decline in marketized public services: An experimental evaluation of choice overload. *Journal of Public Administration Research and Theory*, 26(3), 421-432.

- Johannesen, R. L. (1974). The Functions of Silence: A plea for communication research. *Western Speech*, 38(1), 25–35.
- Kelley, H. H., & Michela, J. (1980). Attribution theory and research. *Annual Review of Psychology*, 31, 457–501.
- Kelley, H. H., & Michela, J. (1980). Attribution theory and research. *Annual Review of Psychology*, 31, 457–501.
- Krause, G. A., & Douglas, J. W. (2005). Institutional design versus reputational effects on bureaucratic performance: Evidence from U.S. government macroeconomic and fiscal projections. *Journal of Public Administration Research and Theory*, 15(2), 281–306.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480-498.
- Larson, A. (1992). Network dyads in entrepreneurial settings: A study of the governance of exchange processes. *Administrative Science Quarterly*, 37, 76-104.
- Lavertu, S. & Moynihan, D. P. (2012), Agency political ideology and reform implementation: Performance management in the Bush administration. *Journal of Public Administration Research and Theory*, 23(3), 521-549.
- Lee, D., & Van Ryzin, G. G. (2019). Measuring bureaucratic reputation: Scale development and validation. *Governance*, 32(1), 177-192.
- Lee, D., & Van Ryzin, G. G. (2020). Bureaucratic reputation in the eyes of citizens: an analysis of US federal agencies. *International Review of Administrative Sciences*, 86 (1), 183-200
- Lee, S. Y., & Whitford, A. B. (2013). Assessing the effects of organizational resources on public agency performance: Evidence from the US federal government. *Journal of Public Administration Research and Theory*, 23(3), 687-712.
- Lowe, G. D. & Pinhey, T. K. (1982). Rural-urban differences in support for environmental protection. *Rural Sociology*, 47(1), 114.
- Luoma-aho, V. (2008). Sector reputation and public organisations. *International Journal of Public Sector Management*, 21(5), 446-67.
- Maor, M. (2007). A scientific standard and an agency’s legal independence: Which of these reputation protection mechanisms is less susceptible to political moves? *Public Administration*, 85(4), 961–978.
- Maor, M. (2011). Organizational reputations and the observability of public warnings in 10 pharmaceutical markets.” *Governance*, 24 (3):557–82.
- Maor, M. (2015). Theorizing bureaucratic reputation. In Maor, M & Waeraas, A. (Eds.), *Organizational reputation in the public sector* NY: New York: Routledge, 18-36.

- Maor, M. (2016). Missing areas in the bureaucratic reputation framework. *Politics and Governance*, 4(2), 80.
- Maor, M., & Sulitzeanu-Kenan, R. (2013). The effect of salient reputational threats on the pace of FDA enforcement. *Governance*, 26(1), 31-61.
- Maor, M., & Sulitzeanu-Kenan, R. (2016). Responsive change: Agency output response to reputational threats. *Journal of Public Administration Research and Theory*, 26(1), 31-44.
- Maor, M., Gilad, S., & Bloom, P. B. N. (2013). Organizational reputation, regulatory talk, and strategic silence. *Journal of Public Administration Research and Theory*, 23(3), 581–608.
- Mariconda, S., & Lurati, F. (2014). Being known: A literature review on media visibility, public prominence and familiarity with implications for reputation research and management. *Corporate Reputation Review*, 17(3), 219–236.
- Marvel, J. D., & Girth, A. (2015). Citizen attributions of blame in third-party governance. *Public Administration Review*, 76(1), 96-108.
- Matheson, G. and Wearing, M. (1999). Within and without: Labour force status and political views in four welfare states. *The End of the Welfare State*: 135-160.
- McAuley, E., Duncan, T., & Russell, D. (1992). Measuring causal attributions: The revised causal dimension scale (CDSII). *Personality and Social Psychology Bulletin*, 18(5), 566–573.
- McGraw, K. M. (1990). Avoiding blame: An experimental investigation of political excuses and justifications. *British Journal of Political Science*, 20(1), 119-131.
- McGraw, K. M. (1991). Managing blame: An experimental test of the effects of political accounts. *The American Political Science Review*, 85(4), 1133-1157.
- McGuire, William J. (1985). Attitudes and Attitude Change, in G. Lindzey and E. Aronson (Eds.). *The Handbook of Social Psychology*. Vol. 2. New York, NY: Random House.
- McKnight, H. D., Choudbury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research*, 13(3), 334–359.
- Meier, K. (1987). *Politics and the bureaucracy: Policymaking in the fourth branch of government* (2nd edition). California: Brooks/Cole Publishing Company.
- Meier, K. J. & Nigro, L. G. (1976). Representative bureaucracy and policy preferences: A study in the attitudes of federal executives. *Public Administration Review*, 36(4), 458-69.

- Meier, K. J. (1992). *Politics and the bureaucracy: Policymaking in the fourth branch of government*. Pacific Grove, CA: Brooks/Cole Publishing Company.
- Meier, K. J., & Bothe, J. (2007). *Politics and the Bureaucracy* (5th ed.). Boston, MA: Cengage Learning.
- Milward, B. H., & Provan, K. (2000). Governing the hollow state. *Journal of Public Administration Research and Theory*, 10(2), 359–80.
- Miller, G.J. and Whitford, A.B. (2016). *Above politics: Bureaucratic discretion and credible commitment*. UK, Cambridge: Cambridge University Press.
- Moe, T. M. (1984). The new economics of organization. *American Journal of Political Science*, 28(4), 739-777.
- Moore, C. D. (2015). Innovation without reputation: How bureaucrats saved the veterans' Health Care System." *Perspectives on Politics*, 13 (2), 327–44.
- Morgeson, F. V., & Petrescu, C. (2011). Do they all perform alike? An examination of perceived performance, citizen satisfaction and trust with US federal agencies. *International Review of Administrative Sciences*, 77(3), 451–479.
- Mortensen, P. B. (2012). "It's the central government's fault": Elected regional officials' use of blame-shifting rhetoric. *Governance*, 25(3), 439-461.
- Moynihan, D. P. (2012). Extra-Network organizational reputation and blame avoidance in networks: The Hurricane Katrina example. *Governance*, 25(4), 567–588.
- Nielsen, P. A., & Baekgaard, M. (2015). Performance information, blame avoidance, and politicians' attitudes to spending and reform: Evidence from an experiment. *Journal of Public Administration Research and Theory*, 25(2), 545-569.
- O'Toole, L. J. (1997). Treating networks seriously: Practical and research-based agendas in public administration. *Public Administration Review*, 57, 45–52.
- Osborne, D., & Gaebler, T. (1993). *Reinventing government: How the entrepreneurial spirit is transforming government*. New York, NY : Penguin.
- Ostrom, E. (1998). A behavioral approach to the rational choice theory of collective action. *The American Political Science Review*, 1, 1-22.
- Overman, S., Busuioc, M., & Wood, M. (2020). A multidimensional reputation barometer for public agencies: A validated instrument. *Public Administration Review*, <https://doi.org/10.1111/puar.13158>.
- Page, B. I. & Shapiro, R. Y. (2010). *The rational public: Fifty years of trends in Americans' policy preferences*. Chicago: University of Chicago Press.
- Perry, J. L. (1997). Antecedents of public service motivation. *Journal of Public Administration Research and Theory*, 7(2), 181-97.

- Peters, G. B., & Pierre, J. (1998). Governance without government? Rethinking public administration. *Journal of Public Administration Research and Theory*, 8(2), 223-43.
- Pew Research Center. (2015). Public's policy priorities reflect changing conditions at home and abroad. Retrieved from <http://www.people-press.org/2015/01/15>
- Pew Research Center. Oct. 2013. "IRS Viewed Least Favorably among Federal Agencies." www.pewresearch.org/fact-tank/2013/10/21/irs-viewed-least-favorably-among-federal-agencies/.
- Picci, L. (2011). *Reputation-based governance*. Stanford, CA: Stanford University Press.
- Pollit, C. (2011). Performance blight and the tyranny of light? Accountability in advanced performance measurement regimes. In M. J. Dubnick, & G. Frederickson, *Accountable governance: Problems and promises* (pp. 81-97). New York, NY: M. E. Sharpe.
- Rainey, H. (2009). *Understanding and managing public organizations* (4th ed.). San Francisco, CA: Jossey-Bass.
- Ray, S. J. (1999). *Strategic communication in crisis management lessons from the airline industry*. Westport, CT: Quorum.
- Remler, D. K., & Van Ryzin, G. .. (2015). *Research methods in practice: Strategies for description and causation*. (2nd ed.). Thousand Oaks, CA: SAGE Publications.
- Rhodes, R. A. (1996). The new Governance: Governing without government. *Political Studies*, XLIV, 652-667.
- Rice, T. W., McLean, W. P. and Larsen, A. J. (2002). Southern distinctiveness over time, 1972-2000. *American Review of Politics*, 23, 193-220.
- Ricketts, T. C. (2000). The changing nature of rural health care. *Annual Review of Public Health*, 21(1), 639-657.
- Roberts, P. S. (2006). FEMA and the prospects for reputation-based autonomy. *Studies in American Political Development*, 20 (1), 57-87.
- Roberts, P. W., & Dowling, G. R. (2002). Corporate reputation and sustained financial performance. *Strategic Management Journal*, 23, 1077-1093.
- Rohr, J. A. (1988). Bureaucratic morality in the united states. *International Political Science Review*, 9(3), 167-78.
- Romzek, B. S., & Dubnick, M. J. (1987). Accountability in the public sector: Lessons from the challenger tragedy. *Public Administration Review*, 47(3), 227-238.

- Rosenau, J. N. (1992). Governance, order and change in world politics. In J. N. Rosenau, E.-O. Czempiel, & S. Smith, *Governance without government: Order and change in world politics* (pp. 1-29). Cambridge, UK: Cambridge University Press.
- Rudolph, T. J., & Evans, J. (2005). Political trust, ideology, and public support for government spending. *American Journal of Political Science*, 49(3), 660–671.
- Sarstedt, M., & Schloderer, M. P. (2010). Developing a measurement approach for reputation of non-profit organizations. *International Journal of Nonprofit and Voluntary Sector Marketing*, 15(3), 276-99.
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online*, 8(2), 23-74.
- Schwaiger, M. (2004). Components and parameters of corporate reputation - An empirical study. *Schmalenbach Business Review*, Vol. 56(January), 46–71.
- Seltman, H. J. (2014). Experimental design and analysis. Retrieved from http://www.stat.cmu.edu/_hseltman/309/Book/Book.pdf
- Shamsie, J. (2003). The context of dominance: An industry-driven framework for exploiting reputation. *Strategic Management Journal*, 24, 199-215.
- Shapiro, R. Y. & Mahajan, H. (1986). Gender differences in policy preferences: A summary of trends from the 1960s to the 1980s. *Public Opinion Quarterly*, 50(1), 42-61.
- Sievert, M., Vogel, D., Reinders, T., & Ahmed, W. (2020). The power of conformity in citizens' blame: Evidence from a survey experiment. *Public Performance & Management Review*, 43(1): 53-80
- Steffel, M., Williams, E., & Permann-Graham, J. (2016). Passing the buck: Delegating choices to others to avoid responsibility and blame. *Organizational Behavior and Human Decision Processes*, 135, 32-44.
- Stoker, G. (1998). Governance as theory: Five propositions. *International Social Science Journal*, 50(155), 17-28.
- Sulitzeanu-kenan, R., & Zohlnhofer, R. (2019). Policy and blame attribution: Citizens' preferences, policy reputations, and policy surprises. *Political Behavior*, 41(1), 53-77.
- Thompson, G., Frances, J., Levacic, R., & Mitchell, J. (1991). *Markets, Hierarchies and Networks*. London: Sage.
- Titchener, E. B. (1910). *Textbook of psychology*. New York: Macmillan.
- Van de Walle, S. (2016). When public services fail: A research agenda on public service failure. *Journal of Service Management*, 27(5), 831-846.

- Van De Walle, S., & Van Ryzin, G. G. (2011). The order of questions in a survey on citizen satisfaction with public services: Lessons from a split-ballot experiment. *Public Administration*, 89(4), 1436–1450.
- Van de Walle, S., Van Roosbroek, S., & Bouckaert, G. (2008). Trust in the public sector: is there any evidence for a long-term decline? *International Review of Administrative Sciences*, 74(1), 47-64.
- van Engen, N. A. (2017). A short measure of general policy alienation: Scale development using a 10-step procedure. *Public Administration*, 95(2), 512-526.
- Van Ryzin, G. G. (2004). Expectations, performance, and citizen satisfaction with urban services. *Journal of Policy Analysis and Management*, 23(3), 433-448.
- Van Ryzin, G. G. (2013). An experimental test of the expectancy-disconfirmation theory of citizen satisfaction. *Journal of Policy Analysis and Management*, 32(3), 597-614.
- Van Ryzin, Gregg G. (2011). Outcomes, process, and trust of civil servants. *Journal of Public Administration Research and Theory*, 21(4), 745-760.
- Wæraas, A., & Byrkjeflot, H. (2012). Public sector organizations and reputation management: Five problems. *International Public Management Journal*, 15(2), 186–206.
- Walker, K. (2010). A systematic review of the corporate reputation literature: Definition, measurement, and theory. *Corporate Reputation Review*, 12(4): 357-387.
- Walsh, G., and Beatty, S. E. (2007). Customer-based corporate reputation of a service firm: Scale development and validation. *Journal of the Academy of Marketing Science*, 35(1): 127-43.
- Wamsley, G. L., & Zald, M. N. (1973). *The political economy of public organizations: A critique and approach to the study of public administration*. Lexington, MA: Lexington Books.
- Watkins, J. (1952). The principle of methodological individualism. *The British Journal for the Philosophy of Science*, 3, 186-189.
- Weber, M. (1922). *Economy and Society*. In G. Roth, & C. Wittich. Berkeley, CA: University of California Press, 1986.
- Weiner, B. (1980). *Human motivation*. New York, NY: Holt, Rinehart and Winston.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92(4), 548-573.
- Weiner, B. (2000). Attributional thoughts about consumer behavior. *Journal of Consumer Research*, 27(3), 382-387.

- Weiner, B., Perry, R., & Magnusson, J. (1988). An attributional analysis of reactions to stigmas. *Journal of personality and social psychology*, 55(5), 783-748.
- Willbern, Y. (1984). Types and levels of public morality. *Public Administration Review*, 44(2), 102-8.
- Williamson, O. E. (1999). Strategy research: Governance and competence perspectives. *Strategic Management Journal*, 20, :1087-108.
- Wilson, S. R., Cruz, M., Marshall, L., & Rao, N. (1993). An attributional analysis of compliance-gaining interactions. *Communication Monographs*, 60(4), 352–372.
- Worthington, R. L., & Whittaker. T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, 34(6), 806–38.
- Wright, B. E., Christensen, R. K., & Pandey, S. K. (2013). Measuring public service motivation: Exploring the equivalence of existing global measures. *International Public Management Journal*, 16(2), 197-223.
- Yang, S. U. (2007). An integrated model for organization—public relational outcomes, organizational reputation, and their antecedents. *Journal of Public Relations Research*, 1 (2), 91–121.

APPENDIX A.

Description of the Sample-Chapter 2.

Characteristics		Sample (%)
gender	Male	34.29
	Female	65.71
race	White	84.10
	Black	5.49
	Hispanic	3.47
	Asian	4.62
	Others	2.31
age	18-24	0.87
	25-34	13.08
	35-44	20.35
	45-54	20.64
	55 or older	45.06
education	Lesser than High school	2.04
	High school graduate	36.44
	Two-year college	12.83
	Four-year college	30.61
	Graduate school	18.08
income	< 29,999	23.75
	30,000-79,999	46.63
	80,000-150,000	22.87
	More than 150,000	6.74
location	A large city	34.97
	A suburban area	32.08
	A small city	19.94
	rural area	13.01
political ideology	Left	29.86
	Middle	24.05
	Right	46.09
		N=341=347

APPENDIX B.

Exploratory Factor Analyses of Split-Half Samples-Chapter 2.

Variable	Items	Split-half samples 1-5 (Loadings on Factor 1 only)				
		1	2	3	4	5
item1	I have a favorable opinion about this agency	0.61	0.62	0.59	0.62	0.53
item2	I believe this agency is doing a good job	0.69	0.69	0.67	0.65	0.53
item3	I don't have much respect for this agency					
item4	Overall, this agency has a good reputation	0.56	0.66	0.62	0.58	0.51
item5	This agency is a waste of taxpayer's money					
item6	I have a negative impression about this agency					
item7	This agency is a well-run organization	0.69	0.66	0.64	0.62	0.50
item8	This agency is effective at its job	0.51	0.60	0.51	0.53	0.47
item9	This agency does a poor job					
item10	This agency is a high performing agency	0.59	0.63	0.62	0.57	0.44
item11	This agency often fails to get things done					
item12	This agency has the capacity to get things done	0.47			0.52	0.51
item13	This agency can be trusted to do what is right	0.68	0.50	0.48	0.68	0.55
item14	This agency maintains high ethical standards	0.68	0.44	0.51	0.66	0.62
item15	This agency seems to be corrupt					
item16	This agency protects democratic values	0.68	0.47	0.53	0.69	0.5
item17	This agency is sometimes dishonest					
item18	I believe what this agency says	0.73	0.74	0.68	0.70	0.69
item19	This agency respects due process	0.78	0.81	0.77	0.73	0.76
item20	This agency is highly transparent	0.82	0.60	0.67	0.83	0.85
item21	Although I sometimes disagree with its decisions, this agency always follows the rules	0.61	0.73	0.73	0.64	0.64
item22	This agency treats people fairly	0.80	0.77	0.77	0.79	0.78
item23	This agency protects the rights of citizens	0.78	0.77	0.77	0.74	0.75
item24	This agency is politically neutral	0.84	0.56	0.59	0.81	0.81
item25	This agency has the technical expertise to do its job well		0.40			
item26	This agency is technically competent					
item27	This agency has highly skilled employees				0.42	
item28	This agency bases its decisions on evidence	0.77	0.76	0.67	0.61	0.65
item29	This agency has the skill to deal with complex situations	0.44	0.64	0.54	0.50	0.44
item30	This agency lacks technical knowledge					
Split-half sample size n		167	167	166	166	163

Note: Principal components factoring (pcf) in Stata with varimax rotation run on five random split-half samples. Only loadings on the dominant first factor are shown. Selected scale items and related loadings are bolded.

APPENDIX C.

Bureaucratic Reputation Scale-Chapter 2.

Bureaucratic Reputation Scale (BRS)

Please indicate the extent to which you agree, or disagree with the following statement about [NAME OF AGENCY].

Dimensions	Items	Strongl y disagre e	Disagre e	Neither agree nor disagree	Agree	Strongl y agree
Short (5-item) scale						
General	Overall, this agency has a good reputation	1	2	3	4	5
Performanc e	This agency is a well-run	1	2	3	4	5
Moral	This agency maintains high ethical standards	1	2	3	4	5
Procedure	This agency treats people fairly	1	2	3	4	5
Technical	This agency bases its decisions on evidence	1	2	3	4	5
Additional items for long (10-item) scale						
General	I have a favorable opinion about this agency	1	2	3	4	5
Performanc e	This agency is a high performing agency	1	2	3	4	5
Moral	I believe what this agency says	1	2	3	4	5
Procedure	This agency is politically neutral	1	2	3	4	5
Technical	This agency has the skill to deal with complex situations	1	2	3	4	5

Note: When possible, randomize the order of items.

