



The Performance Information Processing Framework: Four Cognitive Models of Performance Information Use

Sean Webeck

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

June 10, 2019

The Performance Information Processing Framework: Four cognitive models of performance information use

Sean Webeck

Defense Resources Management Institute

Naval Postgraduate School

*This manuscript is in progress and will be presented at PMRC 2019.
Please do not cite without permission from the author.*

Abstract The question of how public managers use public sector performance information received significant scholarly attention in recent years. The promise of performance management systems was to rationalize the decision making process by creating objective performance metrics that citizens, political officials, and public managers could use to assess the performance of public organizations. Some theoretical work suggests, however, that there is a certain subjectivity to these data, which arises from an individual's role in their organization or broader political environment. Furthermore, a recent spate of experimental work in this area suggests subjectivity might also arise, at the individual level, through cognitive bias. I bridge these two bodies of scholarship with a framework of performance information processing, which incorporates four models of political information use into the story of how public managers use performance information. I suggest that cognitive bias can contribute to the subjectivity of performance information when public managers *process* performance information. In other words, a model of *meaning avoidance* suggests that managers accurately receive performance information from management systems, but that cognitive biases influence the ways in which they interpret or act upon that information. In this essay, I provide empirical evidence for this model. I show that despite different presentations, public managers can accurately recount the objective information they saw when asked to recall it. I also provide evidence that despite being equally aware of objective raw performance metrics, public managers exhibit evidence of cognitive bias when asked to interpret the meaning of that information. This study contributes to the broader discussion of how individuals use performance information.

INTRODUCTION

“Empiricism assumes that objects can be understood independently of observing subjects. Truth is therefore assumed to lie in a world external to the observer whose job is to record and faithfully reflect the attributes of objects.” (Harvey, 2001)

“Against positivism, which halts at phenomena—“there are only facts”—I would say: no, facts is precisely what there are not, only interpretations.” Friedrich Nietzsche, as cited in (Cox, 1999)

How individuals use the performance information created by now ubiquitous performance management systems is one of the big questions facing public management scholars today. One of the challenges in tackling this question is our ability to adequately model the broader set of actions and processes that ultimately contribute to performance information use. The purpose of this paper is to present a framework based on cognitive psychology that describes four different models of performance information use along with empirical evidence to support one of these models.

In 1995, Robert Behn suggested that one of the big questions facing scholars who study public sector organizations was understanding how *“public managers use measures of the achievements of public agencies to produce even greater achievements”* (Behn, 1995). Yet, recently, 20 years after Behn pointed out a major question for researchers in the field, a prominent scholar commented that *“we know little about the basic tendency of individuals to incorporate and use performance information”* (Moynihan, 2015). While scholars have developed frameworks, such as the Interactive Dialogue Model (IDM), that significantly contribute to our understanding of performance information (Moynihan, 2008), there is more to learn. For example, in the decade since the IDM’s initial publication, we have developed several key insights into the phenomenon of interest. But, as a community, we have yet to adequately update our dominant models.

More recently, a new research program on the behavioral foundations of performance information use has contributed a rash of empirical evidence for how individuals process performance information. This behavioral turn has yielded important insights because, typically, scholarship on how public managers use performance information had relied on survey responses and self-reported information (Ammons & Rivenbark, 2008; Kroll, 2015a; Donald P. Moynihan, Poul A. Nielsen, & Alexander Kroll, 2017).¹ Behavioral research often uses different methodologies and theoretical foundations from research based in organizational theory, allowing for different insights. Following this behavioral approach, I aim to contribute to the understanding of how public managers use performance information in two ways. First, I argue the information processing approach allows us to gain new leverage on the question of performance information use because it requires us to look at individual steps in the way that an individual *processes*, or make decisions about, information (Oppenheimer & Kelso, 2015). Second, I suggest the importance of the role of the *interpretation* of performance information as a cognitive step that creates subjectivity in the use of objective information generated by performance management systems (Gaines, Kuklinski, Quirk, Peyton, & Verkuilen, 2007).

Insights into the psychological factors that influence how public managers process performance information may have implications for performance management. On the one hand, how public managers *interpret* performance information potentially influences how public managers “use” performance information. Current scholarship has not paid adequate attention to this important antecedent. On the other hand, an information processing approach, rooted in psychology, also suggests important limitations in dominant models of performance information use, such as the Interactive Dialogue Model (IDM). Specifically, while current models allow for

¹ Typical approaches include single-city case studies, multicity surveys, and multicity case studies.

certain types of subjectivity—e.g., those arising from organizational factors—they cannot adequately incorporate well established cognitive biases in human decision making.

I borrow a framework from political psychology on political information use to address this shortcoming (Gaines et al., 2007). The Gaines et al. framework of political information use consists of four cognitive processing models that seek to describe how individuals process information over multiple steps. Here, I offer a framework of how individuals process *performance information*, offering “awareness” and “interpretation” as distinct steps in how people process this type of information. I believe this will improve our ability to describe and predict. In short, this approach can help scholars interested in how individuals use performance information to develop theory in this area.

As I will show in this essay, while individuals tend to be equally aware of the objective value of the performance information they observe, their interpretations are prone to cognitive biases that arise in large part from the way in which information is presented to them. That is, subjectivity arises specifically from the action and process of *interpretation* and not directly from the information itself nor in how individuals initially incorporate new information. Taken together, this model and relevant empirical findings allow for a better understanding of how the design of performance measurement systems can influence performance information use. Since empirical findings from the broader study of performance information use suggest there is value in updating the IDM, this framework can help students of public management begin to synthesize prevailing models with evidence from how individuals—citizens, politicians, or public managers—use performance information.

In this essay I empirically test the claim that individuals are objectively aware of the performance information they observe but the process of interpretation leads to subjectivity in performance information. This allows me to demonstrate the validity of the larger claim that it is the process of interpreting performance information, and not how individuals receive performance information, that is subject to cognitive bias.

In what follows, I review scholarly work on performance information use. I emphasize the role of institutions and psychology in this research program. I also review and provide a discussion of Moynihan's (2008) Interactive Dialogue Model, including its key assumptions. Subsequent empirical work suggests some potential ways to update the model. For example, I suggest how an understanding of human behavior, as well as recent experimental evidence on the study of performance information use, provide support for incorporating an information processing perspective into the IDM. Then, I point to work from Gaines and colleagues (2007) to potentially address these challenges and refine the larger IDM. I follow with a set of general propositions and specific hypotheses for ways in which cognitive biases can influence how individuals process performance information. I then provide empirical evidence for this framework. I conclude with a discussion of these results and by highlighting avenues of future research.

LITERATURE REVIEW

Performance Management Systems

When it comes to the public sector, performance management regimes are everywhere. As one scholar noted, "the dissemination of quantitative measures of performance has been one of the most widespread trends in government in past decades" (Moynihan, 2015, 33). And, throughout the performance management movement, leading scholars have tried to direct attention to the

question of how public managers use performance information (Behn, 1995; D. P. Moynihan & S. K. Pandey, 2010). Yet, recently, a prominent scholar commented that “we know little about the basic tendency of individuals to incorporate and use performance information” (Moynihan, 2015, p. 33). If we endeavor to learn, and if we are to accept the claim that how public managers use performance information is an important question for public management researchers (D. P. Moynihan & S. K. Pandey, 2010), we must acknowledge some of the limitations of previous (Kroll, 2015a; Donald P. Moynihan et al., 2017).

Institutions and Performance Information Use

Early studies on public managers focused on the question of how aspects of the institutional environment drive performance information use. A recent systematic literature review of the use of performance information outlined some of the progress made in this area since 2000 (Kroll, 2015a). In that time, in “a highly relevant and fast-growing research area” (ibid., 460), research consistently shows six factors commonly drive the use of performance information among bureaucrats: measurement system maturity (for examples see Berman & Wang, 2000; Ho, 2006; Taylor, 2009), stakeholder involvement (for examples see Bourdeaux & Chikoto, 2008; Ho, 2006; Donald P Moynihan & Sanjay K Pandey, 2010), leadership support (for examples see Moynihan & Ingraham, 2004; Moynihan & Lavertu, 2012; Yang & Hsieh, 2007), support capacity (for examples see Berman & Wang, 2000; Julnes & Holzer, 2001; Moynihan & Hawes, 2012), innovative culture (for examples see Moynihan, 2005; Donald P Moynihan & Sanjay K Pandey, 2010; Moynihan, Pandey, & Wright, 2012b), and goal clarity (for examples see Moynihan & Landuyt, 2009; Moynihan, Pandey, & Wright, 2012a; Moynihan et al., 2012b). Thus, previous scholarship suggests that, as expected by institutional theory, the involvement of external

stakeholders influences if and how bureaucrats use performance information. Relevant to this study, a majority of previous studies focus on the ways in which the organizational context influences performance information use.

An Interactive Dialogue Model

In 2008, Moynihan laid out what he referred to as the “Interactive Dialogue Model” (IDM). This model describes how and why a public manager’s institutional context influences her use of performance information. Thus, the pre-eminent framework for understanding performance information use demonstrates the importance of institutions in understanding the subject.

The foundation of this model is a view that performance management systems are decision making systems and performance management is a decision making problem. He defines performance management as “a system that generates performance information through strategic planning and performance measurement routines and that connects this information to decision venues, where, ideally, the information influences a range of possible decisions” (2008, p. 5). In this view of performance management systems, to understand how public managers use performance information, we must understand how they *make decisions*, about performance (information).

While scholars have presented different takes on what it means to “use” performance information—see, for example, Behn (1995)—Moynihan (2008) suggests the purpose of use is ultimately to persuade. In line with the idea of persuasion as the aim of use, at the time of its publication, the most important contribution of the model was the idea that performance information could be subjective. For example, one essential element of the IDM is that

performance information can be subjective through deliberate persuasion. That is, public managers *knowingly* add subjectivity to performance information.

Even more acutely, performance information can be ambiguous because of political considerations *a priori* to any descriptive story of performance information “use”. This idea of subjectivity is contrary to the performance management doctrine, representing a major break in the theoretical development of performance information use by public managers (Moynihan, 2008). This last point helps us understand why the IDM came to provide the theoretical backdrop for a decade’s worth of research. But, as I will discuss later, subsequent research has demonstrated some potential limitations to this perspective

Elements of the Model

There are three fundamental elements of the model. They are: (1) performance information, (2) the individual decision maker (i.e., public managers), and (3) the environment(s) in which these other elements exist or operate. Here, environment is meant to imply both the organizations in which individuals work and the political environment(s) in which those individuals and organizations are situated. The model’s emphasis on organizational and environmental factors (see Moynihan 2008, p. 103) parallels many other studies in the research program (Kroll, 2015a).

Taken together, these three elements shaped the way scholars studied performance information use over the last decade. Grounded in the logic of institutions, the IDM gives us a story in which institutions matter. Yet, the emphasis on institutions raises the question of what other elements might influence and help explain how public managers use performance information. One potential avenue of explanation is the relationship between performance information use and factors at the level of the individual (Kroll, 2015a).

Assumptions of the model

There are six basic assumptions to the model (Moynihan, 2008, 102). First, performance information is not comprehensive. Second, performance information is ambiguous. Third, performance information is subjective. Fourth, the production of performance information does not guarantee use. Fifth, institutional affiliation and individual beliefs will affect selection, perception, and presentation of performance information. Sixth, the concept of dialogue will affect the ability to use performance information to develop solutions.

Flow of the Model

In this model, organizational and political factors influence performance information (use) in several ways. This includes, the presentation of performance information, whether an individual considers (i.e., “looks at”) performance information, how they interpret performance information, and, finally, how they “use” performance information. For anyone interested in studying how people “use” information, this seems like it could oversimplify how individuals interact with information.

To be clear, the IDM assumes performance information is subjective because: (1) individuals can choose to present information subjectively, (2) even the act of considering performance information is, in and of itself, a choice, (3) individuals interpret performance information based upon organizational and political factors, and (4) individuals use performance information to strategically achieve organizational and personal objectives. These assumptions lead to a meta-assumption: “that simply because performance information exists, there is no

guarantee that it is used” (Moynihan, 2008, 102). I want to push on this meta-assumption because I think it oversimplifies how human beings process information in two key ways.

First, I think it treats use as binary—either people use it, or they do not. Second, here, use is implied to be a discrete act. This idea of use is in line with early writings from Behn (1995). In this sense, use is an action taken whereupon said action has been informed by the performance information in question (Van Dooren & Van de Walle, 2011). In other words, to be adequately considered as performance information “use”, a decision maker must have looked at the performance data in question and both 1) become aware of the performance information and 2) updated (or not²) their interpretation of a policy area. The key assumption here is that once an individual decides to look at performance information that she and all others will incorporate and interpret that information. On one hand, characterizing use in this way facilitates observation. On the other, it does not seem to hold up to some basic assumptions from information processing theory (Oppenheimer & Kelso, 2015). To clarify this last point, existing work makes critical—yet unstated—assumptions about how individuals process performance information. In fact, there is growing evidence of the value of incorporating an information processing approach when considering the subject of decision making. This general approach suggests *basic models of cognition should form the basis for how we conceptualize human decision making*. These models allow us to focus our attention on “how decision-relevant information is sampled, retrieved, and integrated” (Oppenheimer & Kelso, 2015, p. 283, 283).

Psychology and the Subjectivity of Performance Information

² This part could be flexible in a theoretical sense, but it is left unspecified in the IDM.

Another reason performance information could be subjective can be found in the way that human beings process information. With the recent emphasis on developing a behavioral public administration, the contours of such an approach are already in place (Grimmelikhuijsen, Jilke, Leth Olsen, & Tummers, 2016). Additionally, some scholars have pointed to behavioral factors that might produce systematic variation in the use of performance information among bureaucrats (Kroll, 2015a; Donald P Moynihan, Poul Aaes Nielsen, & Alexander Kroll, 2017).

While the IDM is largely bereft of psychology as an influence in the larger process of using performance information, it points to a potential role for psychology in a confirmation bias. For example, consider the following falsifiable hypothesis: “Different actors can examine the same performance information and come up with competing, though reasonable, arguments for what the information means” (2008, 113). Carrying on with this line of thinking, Moynihan says:

“Performance information does not necessarily result in clearer decisions if the actors involved cannot agree on what it tells them about current performance, changing budgets, or management. As roles motivate the actors involved to understand performance information differently, the inherent ambiguity in performance information will be exploited.” (2008, 16-17).

The preceding statement makes three key assumptions which should be explored in more detail. First, the selection of a piece of performance information is a deliberate (i.e., cognitive) act rising from one’s role. Second, there is a question about whether actors can even agree on what performance information tells them. This may arise from, among other things, uncertainties in the broader political environment. A third assumption is that astute political actors will deliberately exploit potential ambiguities of performance information through interpretation. These are key

assumptions of an important model in this literature. For this reason, I argue public management scholars should give more attention to these assumptions.

In the model, because ambiguity arises through role, there is little space for other factors to contribute to subjectivity. Yet, recent empirical work suggests cognitive processes may also contribute to performance information's ambiguity and subjectivity. The standard take on bounded rationality is that human beings have significant constraints on their ability to hold and process information. In the IDM, however, information overload leads individuals not to "try to process all information but select information that they find useful" rather than simply being unable to accurately process the information they have (Moynihan, 2008, 17).

If we can learn from work on information processing and heuristics, there are several reasons to suggest why individuals respond differently to the same information (Gigerenzer & Gaissmaier, 2011; Oppenheimer & Kelso, 2015). And, many of these different interpretations arise not from deliberate cognitive processing but from a type of cognitive processing that relies on speedy, snap decisions.

Behavioral Foundations of Performance Information Use: Empirical Evidence

Recent empirical work suggests psychology may play a role in helping us understand how public managers use performance information (Andersen & Moynihan, 2016; Kroll, 2015b; Moynihan, 2008, 2015; Poul A Nielsen, 2013; Salge, 2011). Other studies suggest this perspective can contribute to our understanding of how individuals, broadly considered, respond to performance information. These include studies on citizens (Andersen & Hjortskov, 2015; Baekgaard & Serritzlew, 2016; Barrows, Henderson, Peterson, & West, 2016; Hvidman &

Andersen, 2016; Olsen, 2013, 2015a, 2017) and politicians (George, Desmidt, Nielsen, & Baekgaard, 2016; Poul A. Nielsen & Baekgaard, 2015; Poul A Nielsen & Moynihan, 2016; Olsen, 2014). This approach should be understood within a broader context of seeking to understand the psychological foundations of public administration (Grimmelikhuijsen et al., 2016).

For our purposes, there are two important takeaways from these studies. First, experimental methods are a useful approach to develop our understanding of performance information use across a variety of political actors (Anderson & Edwards, 2015; Bouwman & Grimmelikhuijsen, 2016; James, Jilke, & Ryzin, 2017; Jilke, Van de Walle, & Kim, 2016; Donald P. Moynihan et al., 2017). Second, when it comes to performance metrics, these studies suggest that, depending upon the circumstances, individuals exhibit various cognitive biases and utilize several heuristics when responding to performance information. In short, evidence for cognitive bias in the use of performance information by various actors supports the value of taking an information processing approach (Gigerenzer & Gaissmaier, 2011; Oppenheimer & Kelso, 2015).

Motivated Reasoning

One cognitive bias that has received some attention from public management scholars in recent years is motivated reasoning. It is now well established that “motivation may affect reasoning through reliance on a biased set of cognitive processes: strategies for accessing, constructing, and evaluating beliefs” (Kunda, 1990). Epley and Gilovich argue that our motivations potentially influence the way we process information in one of two ways (2016). First, our biases and preferences might lead us to avoid certain information or to emphasize other pieces of information. Second, once we have information, we are free to interpret how we like.

Public management scholars have only recently begun to try and understand the relationship between motivated reasoning and performance information use. One study shows how political motivations influence how political officials prioritize goals in the face of the information presented to them (Christensen, Dahlmann, Mathiasen, Moynihan, & Petersen, 2018). Another set of studies suggest that individuals' motivations influence the way they interpret performance information (Baekgaard, Christensen, Dahlmann, Mathiasen, & Petersen, 2017; Baekgaard & Serritzlew, 2016; James & Van Ryzin, 2016). Unfortunately, none of the aforementioned studies adequately disentangle the question which arises from Epley and Gilovich—namely, do differences in interpretation exist because of deviations in *acquiring* or *interpreting* performance information? While Epley and Gilovich focus on motivated reasoning, scholars could ostensibly apply a wide array of cognitive biases to the question of if and how individuals acquire and interpret performance information.

Present Stumbling Blocks

I categorize our lack of understanding on this issue into three classes of assumptions. The first involves a set of assumptions we make about performance management systems and the information they create. Two that stand out as relevant to the current study are: 1) “Government can and should make more rational decisions”, and 2) “Performance information will improve decisions and can be used to foster accountability” (Moynihan, 2008, 27). The link between performance information and accountability through the quality of decisions made based upon this information rests on a further assumption about the information itself. Performance data were originally thought to allow for an individual to make an objective assessment of how an organization is doing because it is intended to be “systematic” information (Poul Aaes Nielsen, 2013; Radin, 2006). It’s often numerical format provides a “reassuring status of clarity and

objectivity” (Moynihan, 2008, 95). The promise of this story has led us to a point in which performance management systems are ubiquitous. Nonetheless, while performance management doctrine assumes bureaucrats will incorporate the performance information these systems create into their decision making, at this point there is sufficient empirical evidence to refute the basic assumptions listed here.

The second class of assumptions involves the way we conceptualize how individuals respond to information and make decisions. That the study of human decision making has received a lot of scholarly attention is not surprising. Human beings use information to make decisions is one of the most important questions in the social sciences (Giles, 2011). One consequence of this is the ability to broadly define studies of human decision making as falling within one of two camps. The first second focuses on institutions, demonstrating that under many conditions, humans respond in predictable ways to these institutions. The second centers on the assumptions we make about individuals and how they process information.

Third, what does it mean to “use” performance information? Previous research seems to characterize use as a discrete act that follows from observing a performance metric (Behn, 2003; Moynihan, 2010; Van de Walle & Van Dooren, 2011; Van Dooren & Van de Walle, 2011). This is a critical assumption but one potential shortcoming is the way in which what happens at the micro-level is unobservable. An information processing approach gives scholars more theoretical leverage. It allows us to observe how individuals respond to performance information over multiple cognitive steps. For example, it allows us to observe whether or not individuals are actually “aware” of the information in question (and if so, how well) and, if so, how they “interpret” the information (Gaines et al., 2007).

An additional challenge for advancing theory is that progress thus far considers macro- and micro-level processes independent of one another. That is, studies of performance information use by public managers either incorporate an institutional or psychological perspective but not both. This is in line with a critique from Priem and colleagues about management studies, broadly construed. They suggested that “those individual judgment studies that have been performed by management researchers have almost always stayed within one level of analysis” (Priem, Walters, & Li, 2011, 554-555). One assumption advanced here is that to make progress on the basic question for performance information use by public managers, we must figure out how to integrate both the macro- and micro-level processes that work to shape the phenomenon.

Institutions, Behavior, and Performance Information Use

In the context of this discussion, I suggest two limitations to the IDM. First, more recent empirical evidence suggests other cognitive biases might influence the use of performance information. Second, there is an overreliance on role as an institutional variable. If we understand that the fundamental problem in public administration is that bureaucrats make public decisions that have public consequences, we must recognize that citizens and political officials use a multitude of tools to try use to constrain bureaucratic behavior (Bertelli & Lynn, 2006; Kenneth J. Meier & Bohte, 2007). While the IDM appears to offer the flexibility necessary to accommodate revisions in its assumptions of the role of behavior and institutions in how public managers use performance information, the model presently stands in need of a revision that offers a broader array of variables for public management scholars to consider. Here, I offer aspects of behavior—reference points, framing effects, and negativity bias—and institutions—performance benchmarks

and justification requirements—as elements that might influence how public managers interpret performance information.

THEORY

Whether individuals are aware of political facts and how they form political opinions are fundamental questions in political science. Similarly, scholars of public management are interested in how well public managers understand the performance of their organization and if, and how, they *use* information regarding that performance. Additionally, accountability problems lead public management scholars to focus on how managers make public decisions. Ergo, citizens, political officials, public managers and scholars are all interested in the extent to which there is a connection between facts about public sector organizations and the decisions public managers make relating to that information. One approach from political psychology that may help to inform our understanding of how public managers use performance information considers how individuals move from facts about politics to political opinions (Gaines et al., 2007).

Four Models of Political Information Use

Gaines and colleagues (2007) suggest four models of processing political information in which a) objective information exists in the larger political environment, b) individuals become aware of the information, c) individuals must interpret the political information, and d) finally, individuals must arrive at policy positions. Their framework builds off work that seeks to understand “whether people update and what it means to update” (*ibid.*, 958), which arose from a broader discussion over whether citizens were able to objectively update their policy views when new political information entered the environment (Gerber & Green, 1999; Gerber & Green, 1998)

while other work showed evidence of bias in policy views even after new information became publicly available (Bartels, 2002; Taber & Lodge, 2006).

According to Gaines and colleagues (2007), the interpretation of political information plays an important role in moving from political fact to political opinion because interpretation represents the step in the cognitive process where individuals give meaning to political information. In more colloquial terms, motivated reasoning makes partisanship a strong drug that significantly influences the way that individuals look at—i.e., interpret—political information (Kunda, 1990; Taber & Lodge, 2006). But, they did not advance this same expectation about the ability of individuals to acquire new political information and be aware of it. That is, they did not expect that partisanship influences an individual’s ability to accurately update their understanding of political facts.

More recently, the question of information acquisition has interested scholars because of the discussions of “alternative facts” and “fake news” in the broader political environment. And, as suggested by (Epley & Gilovich, 2016), there is some evidence that motivated reasoning does in fact influence how individuals receive and update on *political* facts (Hochschild & Einstein, 2015; Nyhan & Reifler, 2010; Schaffner & Roche, 2016; Yeo, Cacciatore, & Scheufele, 2015). Without a doubt, the role of bias is an important question in the acquisition and awareness of political information. And, scholars of public management should also inquire about the factors that lead to deviations in the acquisition process when considering performance information. Yet, when Gaines et al. (2007) demonstrate empirically that cognitive bias influences how individuals process political information between steps “b” and “c”, they provide evidence that cognitive biases can arise from the broader cognitive process of interpretation rather than *awareness* or information uptake.

I expect awareness after information acquisition to be considerably less susceptible to the influence of cognitive bias for performance information than for political information. This expectation arises because, at least ostensibly, performance information is objective in a way that political information inherently is not; despite the objections about the objectivity of performance information, the original intention of performance management systems was to create measures of performance that accurately and adequately captured the function in question. For example, telling someone 75% of students passed a Math exam in a school or that 86% of residents expressed satisfaction with a city's road maintenance efforts is likely to invoke less partisan or ideological filtering than telling an individual that 97% of scientists agree with the anthropogenic causes of climate change. As such, I expect the model of performance information processing which emphasizes how cognitive bias can influence the *interpretation*, rather than *acquisition* of, performance information (Model 3, discussed below), represents the best model for understanding how individuals (e.g., public managers) process performance information. In the following section, I modify the Gaines et al. framework to apply to performance information processing.

A Framework of Performance Information Processing

If we consider the similarities between political information and performance information—namely, they are information—then the Gaines framework could apply to performance information use (by citizens, political officials, *and* public managers) as well. Ergo, how public managers interpret performance information might have an important effect on the decisions they make.

As it stands, the concept of interpretation plays a role in the IDM when 1) managers choose to consider information or 2) how they might spin performance information to the benefit of their

organization. But, the IDM does not consider how individuals interpret performance information in line with the experimental evidence which suggests the role of cognitive bias in the use of performance information. Thus, it does not consider the role of information processing in how managers *interpret* performance information. Let's suppose a manager wants to consider and then use performance information. What might this process look like?

Table 1 shows an adapted version of the four models of information updating proposed by Gaines and colleagues (2007)—I suggest referring to them individually as performance information processing models or jointly as the Performance Information Processing Framework (PIPF). The models have been adapted to portray four steps of the updating process for performance information use. First, I discuss the different steps and then I discuss the four models which describe different processes of performance information use.

The first step represents the raw performance metric; a manager sees performance information. In the second step the manager becomes aware of the metric; they can recall the information they just saw. In the third step, a manager interprets the information; the manager gives meaning to the performance information and updates their belief about an organization's performance. Finally, the manager decides to undertake some action based upon this information; in the lexicon of this literature, the manager has "used" the information.

Table 1 - Four Models of Performance Information Use								
Model	Model Name	Step 1	Process	Step 2	Process	Step 3	Process	Step 4
1	<i>Complete updating</i>	Performance Information	→	Awareness	→	Interpretation	→	"Use"
2	<i>Fact avoidance</i>	Performance Information	⊥	Awareness	→	Interpretation	→	"Use"
3	<i>Meaning avoidance</i>	Performance Information	→	Awareness	⊥	Interpretation	→	"Use"
4	<i>"Use" disconnect</i>	Performance Information	→	Awareness	→	Interpretation	⊥	"Use"

Here are four models of performance information use along with relevant steps and representations of cognitive processing in the model. The "→" sign signifies a smooth transition from one step to another. The "⊥" sign represents a cognitive process that does not cleanly flow from one step to the other. For example, in Model 2, individuals may see performance information but not become aware of it (i.e., they would not be able to recall what that information was).

In Model 1, *complete updating*, there is a smooth transition from performance information, to awareness, to interpretation, to use. That is, if a manager sees new information which is contrary to her original understanding of the performance of the organization (Kenneth J Meier, Favero, & Zhu, 2015), she becomes aware of it and interprets that information in a way that accurately reflects that information. In this case, we would expect her to recall and then provide an interpretation that matches or is at least very similar to the performance information she saw. The information is then used in a way that reflects smooth cognitive transitions from step one to steps two and then three.

Model 1, *complete understanding*, reflects early thinking on performance information use, which Moynihan refers to as the performance management doctrine (2008). In other words, for many people this would probably be the normatively preferred model of performance information use. But, previously discussed empirical evidence from experimental work suggests some form of deviation from this ideal—in all three groups of users. So, we must look to some other descriptive model to help us understand how public managers process performance information. A model of cognitive processing allows us to pinpoint where, both in a descriptive and a causal sense, the potential sources of variation arise.

In Model 2, *fact avoidance*, managers see performance information but do not update their awareness of what that information was. Gaines and colleagues say some conditions which might lead to fact avoidance are “willful or accidental ignorance” or “if changing conditions create mental discomfort” (2007, 960). These circumstances might lead people to simply pay less attention to reports of performance changes. In the case of performance management, this might be when a public manager is deeply invested in a project that is being evaluated negatively or if they get information that is drastically different than their worldview. The examples I investigate in this project should not lead to those kinds of cognitive conditions or challenges but this could be an area of interest to scholars in the future.

In Model 3, *meaning avoidance*, managers see performance information and become aware of it but do not change their interpretation of the information. That is, the way they interpret the information does not flow smoothly from the newly acquired information itself. Gaines et al. (2007) illustrated this model to demonstrate the role of partisanship in how Americans interpreted the Iraq War. Even though co-partisans representing the two predominant U.S. political parties were aware of the increase in troop casualties, Republican co-partisans (with a Republican as the sitting-President) interpreted the increase in casualties to be less severe than Democratic co-partisans. They suggest that motivated reasoning or reference biases might represent cognitive biases that lead to meaning avoidance. I also argue there are a wide number of cognitive biases that might affect the way that public managers interpret performance information.

In Model 4, “*Use disconnect*”, managers would become aware and then interpret the performance information but would ultimately not use it in a way that flows smoothly from the interpretation stage. This could happen for a variety of reasons—some individual, others institutional. I presume this is the model that best characterizes the IDM.

General Behavioral Expectations

In the Interactive Dialogue Model, subjectivity arises from a cognitive bias—motivated reasoning—that arises from a public manager’s role within a public sector organization. Here, I argue that public managers may exhibit cognitive biases—e.g., responses to reference points—simply because of the way information is presented to them. I also offer that certain accountability tools—performance benchmarks and decision justification requirements—should moderate these biases under appropriate conditions. In short, public managers may interpret ostensibly objective performance information in a subjective way that simply reflects the use of different points of reference.

The labels “System 1” and “System 2” are used to describe two very different cognitive processes (Kahneman, 2011). System 1 processing reflects an unconscious action where human beings use heuristics to think fast, make many associations, and generally process as much information as possible. System 2 processing, on the other hand, is more deliberate, more “rational”. In this mode, individuals undertake more time and effort to consciously undertake a more reliable decision making process. While the IDM focuses on motivated reasoning, other research on the use of performance information suggests that cognitive bias can be observed when individuals exhibit System 1 processing.

In a broad sense, I expect that when presented with situations that reflect the biases discussed herein, public managers will tend to exhibit System 1 thinking. That is, public managers will respond to reference points in a way that reflects System 1 thinking. I believe these effects will increase in the context of a performance benchmark but decrease when public managers are

asked to provide a justification of their decisions. Evidence for these biases will come when public managers respond to qualitatively equivalent information in substantively different ways. That is, evidence a cognitive bias will be present when public managers *interpret* the same performance information in different ways when the only difference is the way that information is presented to them.

As a caveat, these expectations may depend on the scale of a comparison being made or some other artifact of the interpretative process. That is, the hypothesis in this type of behavioral research is not that everyone will exhibit the cognitive bias in question but that a statistically distinct percentage of individuals will exhibit the bias.

Reference Points and Cognitive Bias

The preceding discussion relies heavily on the idea that reference points will serve as a significant component of the way that public managers will interpret performance information. In the following discussion, I expect reference points will help to facilitate System 1 thinking and cognitive bias because individuals will consider the performance metric in the context of some reference point rather than as an objective measure.

Reference Points

Reference points are “stimuli of known attributes that act as standards against which other categorically similar stimuli of unknown attributes are compared in order to gain information” (Yockey & Kruml, 2009, 97). Reference points represent a significant part of our cognitive processing because our judgment is fundamentally comparative in nature (Mussweiler, 2003). That is due to the fact that our perception is “*reference-dependent*” (Kahneman, 2002, 459,

emphasis in original). And, the way we fixate on reference points tends to add a level of subjectivity to the way we interpret decisions and events around us.

Some argue that public managers will use comparisons in how they think about performance (Kenneth J Meier et al., 2015; Olsen, 2015b). Others provide evidence to suggest of this (Ammons & Rivenbark, 2008; Askim, Johnsen, & Christophersen, 2007; Hammerschmid, Van de Walle, & Stimac, 2013; P. A. Nielsen, 2014). Nonetheless, questions pertaining to the importance of reference points in how bureaucrats use performance information remain largely unexplored.

Institutional Activators of Cognitive Bias

My expectation is that the influence of these well-established cognitive biases on the interpretation of performance information can itself be influenced by institutional (i.e., macro) characteristics. We know that organizations do not present information in a vacuum, but instead require that managers use performance information to make certain comparisons with that information. Very often those comparisons are dictated by previously determined levels of acceptable performance, which are determined by levels of performance in similar organizations. These types of comparison are commonly known as benchmarking. In addition to benchmarking, some performance management systems require managers to justify their interpretation of and reactions to performance information. These accountability systems are meant to ensure productive feedback processes and allow agency leaders and/or political principals to monitor and better understand the use of performance information. This section will briefly discuss the literature on the potential impacts of these institutional features on System 1 vs. System 2 thinking in the use of performance information.

Benchmarking

Simply stated, benchmarking is, “The process of comparing performance across organizations” (Bouckaert & Van Dooren, 2009 156). It comes in two forms (Löffler, 2001). Absolute benchmarking occurs when predefined standards of performance lead to a “pass-fail” approach to understanding organizational performance. In this way, any organization (or individual) can “pass” a benchmark. On the other hand, relative benchmarks require competition between comparable organizations. Here, our focus is on absolute benchmarks.

Justifications

It is well known that while political officials need administrators to undertake actions on their behalf, they are motivated to constrain bureaucratic behavior towards producing a set of decisions and results that are more favorable to their preferences (Seidenfeld, 1996, 2001; William F. West, 1995; William F West, 2004). One tool that is commonly used to constrain bureaucratic behavior is the need to explain—to justify—*ex-post*, the motivations and reasoning behind a particular bureaucratic action. This has been said to affect the underlying psychology of bureaucratic behavior (i.e., decision making) because it “encourages agencies to take greater care when formulating rules, which in turn decreases the likelihood that the rulemaking process will reflect psychological decisionmaking [*sic*] biases” (Seidenfeld, 2001, p. 1060, 1060).

Hypotheses

Pre-Registration

I pre-registered this study and the following expectations with the Evidence in Governance and Politics (EGAP) group, application ID: 20180425AD.

Expectations Arising from the Performance Information Processing Framework

Using the logic of the Performance Information Processing Framework, as well as recent empirical findings on the behavioral foundations of performance information use, I put forward the following general hypothesis: when faced with performance information, public managers will process that information in a way that deviates from the “complete updating” performance information processing model. In the past, systematic deviations from rationality in the assessment of performance information have generally been “the bar” of evidence necessary for claiming that cognitive bias influences the way that individual’s use of performance information. Now, the PIPF allows us to move beyond this limitation and to specify *where* in the cognitive process these biases arise.

Evidence of the role of cognitive bias in the processing of performance information could come in three forms. First, individuals might not adequately update their awareness of the objective information (i.e., *fact avoidance*). Second, if they hold an accurate awareness of what the performance information is, they may simply interpret it in a way that diverges from a smooth transition from step 2 to step 3 (i.e., *meaning avoidance*). Third, if there are no deviations from the original information after individuals interpret it, other factors could lead individuals to use it in a way that does not align with the considered performance information (i.e., “use disconnect”). Finally, role, or other organizational factors, could be one example of a factor that might lead to *use disconnect*. While future research should investigate this step in the Performance Information Processing Framework, for the remainder of this essay the emphasis is on Model 2 and Model 3.

The Gaines et al. framework suggests that Model 3, *meaning avoidance*, will have the best descriptive power of these models. Similarly, I expect that when given performance information, public managers will be able to accurately recall—i.e., they can demonstrate an “awareness” of—previously observed performance information but that these same individuals, faced with the same conditions and performance information of which they are aware, will be prone to exhibit cognitive biases when asked to interpret this information.

H1: When faced with performance information in the context of performance benchmarks and justification requirements, public managers will be able to accurately recall the objective performance information.

Performance Benchmarks

I expect that performance benchmarks will set the reference points that public managers use to determine if current performance is above or below acceptable levels. That is, even though benchmarks often represent a type of institution designed to increase accountability in line with the assumptions of the performance management doctrine, the fact that benchmarks lead individuals to undergo the same type of comparative assessments discussed previously suggests they might contribute to the subjectivity of performance information. For this reason, I expect that performance benchmarks will serve to induce System 1 processing; we should expect to see public managers exhibit cognitive bias in their interpretations of performance information in the presence of performance benchmarks.

H2: Benchmarks will drive respondents to provide higher (lower) performance assessments if the raw metric they see is greater (less) than the benchmark.

Decision Justifications

Finally, in line with Seidenfeld's comments on bureaucratic decision making in the rulemaking process (Seidenfeld, 1996, 2001), I expect the need to justify will induce System 2 thinking. The expectation is that having to justify one's thoughts will lead public managers to undergo a more deliberate thought process when considering performance information. Ergo:

H₃: I expect public managers will be less prone to exhibit psychological biases when they are asked to justify their thought processes.

DATA AND EMPIRICS

Empirical evidence to test these hypotheses come from two related experiments run during the same survey. I ask individuals to interpret performance information and then, at a later point in the survey, respond to a question asking them to name the raw performance metric they saw. In this way, I can manipulate the temporal stages of the Performance Information Processing Framework to test the validity of the *meaning avoidance* model.

Data Collection

I used surveys to collect data for this experiment in three phases. Individuals were paid for their participation in all phases. I designed data collection instruments for each phase using Qualtrics. I utilized TurkPrime (www.turkprime.com) as a third-party platform to collect data from Workers on Amazon's Mechanical Turk (MTurk). TurkPrime offers researchers both greater flexibility and control over the design and implementation of online, crowdsourced research (Litman, Robinson, & Abberbock, 2016).

In the first phase, I ran a short survey that allowed me to screen respondents in two ways. First, respondents were asked to select the sector that best described their primary employment. Possible responses included: private for-profit, private not-for-profit, public, and N/A (e.g., unemployed, out of the workforce, etc.). I provided representative examples in case the sector type would confuse anyone. In addition to this question, I also asked individuals if they had ever worked in each of the three sectors. Respondents could select “yes” or “no” to specific (individual) questions about each sector. If they selected “yes”, respondents then saw an additional question in which they provided a numerical response for the number of years they worked in the respective sector. 5342 unique individuals completed this screening phase. Individuals passed as preliminarily qualified if they indicated they currently worked in the public sector or that they had at least five years of work experience in the public sector. Of these, individuals were disqualified for the following reasons: beeline responses (e.g., people indicated they had worked five years in each sector), 50 or more years of experience in any sector, 60 or more years of combined experience, and anyone who first indicated they worked in the public sector but then later indicated they had never worked in the public sector. 1202 individuals met these qualifications.

I then sent a second survey to these 1202 individuals. This survey included demographic items and scales for the Big 5 personality items and public service motivation. Yet, the real motivation behind this phase was to try and screen out those who passed the first phase of the survey but were not in our population of interest—people with significant public sector work experience; especially, public managers. Someone could easily provide inconsistent answers over time. Individuals might lie in one of the two phases because they believe they know what the researchers are looking for. Or, because multiple individuals use the same MTurk account. I undertook this effort in the hope that I could make a stronger claim about the respondents. I

received 773 responses from this wave. Of these, 479 met the qualifications through both waves of the survey. These were the potential respondents who were notified of the opportunity to undertake this survey experiment.

Experimental Design

Respondents were asked to provide a response to a vignette about business satisfaction rates. In this experiment, I was interested in the role of performance benchmarks and justifications as potential moderators for how individuals process performance information. Table 2 shows the group assignments across these two treatment conditions. Ultimately, the experiment had a 2x2 factorial design with individuals randomly assigned to one of four groups: (Group 1) control, (Group 2) justification, (Group 3) benchmark, and (Group 4) benchmark and justification.

Table 2 - Experiment 1			
Group Assignments by Treatment			
<i>Treatments</i>		<i>Benchmark</i>	
		No	Yes
<i>Justify</i>	No	1	3
	Yes	2	4

All individuals saw this prompt:

For this question, imagine that you are the manager of a business development office for a major metropolitan area. Your city just released its yearly performance metrics and, based on this information, the mayor wants to know how you think the city performed over the course of the last year. For some time, business owners were asked whether they were generally satisfied or generally dissatisfied with the city as a place to do business. In 2011, 57% of business owners indicated they were satisfied with the city as a place to do business. The mayor has tasked your office with improving the business climate in the city.

Individuals in the benchmark groups also saw this sentence at the end of the prompt:

The goal has been to increase the percentage of business owners satisfied with doing business in the city to 67%.

All individuals saw a randomly generated performance rating which indicated that between 62% and 72% of business owners were satisfied with the city as a place of doing business over the past year. Randomizing the observed performance metric is an important component of this test. It allows us to gain a better understanding of how individuals process performance information across a range of potential performance metrics. This also allows us to use a stationary benchmark.

All individuals were asked to assess the performance of the city as a place of doing business for the past year (based on the data they saw). This became the dependent variable for the interpretation experiment. All individuals had an equal probability of seeing a value that was a) less than the benchmark (5/11), b) equal to the benchmark (1/11), or c) greater than the benchmark (5/11). (Note, the first two groups do not see the benchmark.) Half of the respondents also need to justify their performance assessment. All respondents who will justify their responses are told they will have to perform this task before they see the raw performance metric.

Then, at a later point in the survey, individuals were asked to provide the raw performance metric they observed in this experiment. This response became the dependent variable of interest for the awareness experiment.

Results

Awareness

Table 3 provides the means and standard deviation for the value respondents provided for the performance metric they observed. Note, the mean for all respondents was 62.92 but the randomized metric individuals saw had a range of 62-72. I also provide an assessment of the percentage of individuals within each group who listed the exact performance metric they saw. I conducted a one-way ANOVA to determine if the stated value of the observed performance metric (randomized between 62-72) was different across four (a control and three treatment) groups. There was not a statistically significant difference in the reported value of the observed performance metric across these groups as determined by one-way ANOVA ($F(3,350) = 0.56, p = 0.64$). This test passed Bartlett's test for equal variances $\chi^2(3) = 6.2108, p = 0.102$.

Table 3 - Awareness of Observed Satisfaction Rate (Check)				
Treatment Group	Mean	Std. Dev.	N	Pct. with Correct Response
1	62.66	10.13	88	29.55%
2	63.59	12.06	88	43.18%
3	61.86	11.87	92	31.52%
4	63.71	9.71	86	29.07%
Total	62.94	10.99	354	33.33%

In addition to the one-way ANOVA, I also ran an OLS regression that included the dependent variable (value provided), the original raw performance metric, the individual's treatment group, and the outcome score (i.e., the "interpretation" value). This was meant as a harder test. Of note, only the observed satisfaction variable (i.e., the raw performance metric) was statistically significant in the regression model. These results are in Table 4. Taken together, these two tests suggest there is no difference in the level of *awareness* that individuals had based upon randomization. So, any differences between the groups in their *interpretations* must come from that specific step in the PIPF.

Table 4 - OLS Regression	
Variable	Business Satisfaction
Observed Satisfaction	0.74*** (3.90)
Treatment Group	-0.13 (-0.25)
Outcome Score	0.04 (0.90)
Constant	10.54 (0.88)
N	354
Prob > F	0.0001
R-squared	0.0584
Dependent variable is the response to a solicitation for the observed performance metric. *** $p < 0.01$	

Beyond these base checks, I wanted to run a more robust check by comparing all the groups against one another. Those results can be found in Table 5 where I looked at both one-way ANOVA and standard deviation tests for the awareness check across the groups. None of the one-way ANOVA tests showed statistically distinct responses between the groups. Although, one comparison of standard deviation was statistically significant at the $p < 0.05$ level and another at the $p < 0.1$ level. These results appear to be driven by respondents in Group 4 which saw the performance benchmark and the justification requirement. This group had the highest mean but the lowest standard deviation.

Table 5									
Reported Raw Observed Raw Performance Metric (Check): Mean and Variance									
By Treatment Class									
Variable	Group	Mean	Std. Dev.	Std. Err.	ANOVA		SD Test		N
					F	Prob > F	f	Pr (F>f)	
Benchmark	No	63.13	11.11	0.84	0.10	0.7505	1.04	0.7898	176
	Yes	62.75	10.89	0.82					178
Justify	No	62.25	11.03	0.82	1.44	0.2315	1.02	0.9012	180
	Yes	63.65	10.93	0.83					174
By Treatment Group (focus Benchmark)									
Group	Mean	Std. Dev.	Std. Err.	ANOVA		SD Test		N	
				F	Prob > F	f	Pr (F>f)		
1	62.66	10.13	1.08	0.24	0.6279	0.73	0.1365	88	
3	61.86	11.87	1.24					92	
2	63.59	12.06	1.29	0.01	0.9433	1.54	0.0464	88	
4	63.71	9.71	1.05					86	
By Treatment Group (focus Justify)									
Group	Mean	Std. Dev.	Std. Err.	ANOVA		SD Test		N	
				F	Prob > F	f	Pr (F>f)		
1	62.66	10.13	1.08	0.31	0.5795	0.71	0.1060	88	
2	63.59	12.06	1.29					88	
3	61.86	11.87	1.24	1.29	0.2585	1.50	0.0617	92	
4	63.71	9.71	1.05					86	
Control vs Both Treatments									
Group	Mean	Std. Dev.	Std. Err.	ANOVA		SD Test		N	
				F	Prob > F	f	Pr (F>f)		
1	62.66	10.13	1.08	0.49	0.4862	1.09	0.6967	88	
4	63.71	9.71	1.05					86	

Interpretation

Here, I also present evidence on how individuals in this experiment interpreted the performance information they saw. A one-way ANOVA on the assessed organizational performance (i.e., interpretation) showed evidence for a statistically significant relationship for the randomized group assignment ($F(3(350) = 12.31, p = 0.0000)$). Since we know respondents did not show any statistically distinct patterns in their ability to recall performance information *after* they

were asked to interpret it, these results allow us to confidently say that the randomized group assignment influenced how individuals assessed performance through their interpretations of qualitatively similar performance metrics but not in their ability to acquire and be aware of this same information. This is significant evidence in support of Model 3, the *meaning avoidance* performance information model.

Seeing evidence that treatment groups influenced the way individuals assessed performance, the next steps will uncover the reasons for those differences. The three variables of interest are 1) whether the respondent saw a performance benchmark (67% satisfaction), 2) whether the respondent was told they would need to justify their response, and 3) the observed satisfaction value (which was also randomized). The first two assume that the treatment group matters and that it influences how respondents process the observed satisfaction metric. Together, these three allow me to tease out why there are differences in how the groups assessed performance.

Table 7 communicates the most important information about the interpretation experiment. At the treatment level, it compares within both the benchmark and justification treatments. It also compares at the group level, showing which relationships are statistically distinct from one another. All tests of significance in this table were either one-way ANOVA or standard deviation tests, respectively. The importance of the performance benchmark treatment clearly stands out in this table.

Broadly, if I look at each treatment class as a binary variable, both treatments appear to provide statistically significant differences between those who saw the treatment and those who did not. In the benchmark treatment, those who did not see the benchmark had a mean performance assessment of 67.06 while those who saw the performance benchmark provided a mean response

of 73.98. This is a difference of almost seven points and significant at the $p < 0.01$ level. The difference for those who saw the justify treatment was 2.84 points (No, 69.14; Yes, 71.98) and statistically significant at the $p < 0.05$ level. Moving on to the groups, the mean performance assessments for both Group 3 and Group 4 (benchmark groups) were statistically different than their respective non-benchmark comparisons. But, when I look at the justify treatment, it was only significant within the benchmark treatment. That is, without the benchmark, respondents did not interpret performance any differently depending on whether they would need to justify their performance assessment. This finding suggests some limitations to a need to justify performance assessments as a potential moderator in bureaucratic decision making.

Table 7									
Assessed Performance									
<i>By Treatment Class</i>									
Variable	Group	Mean	Std. Dev.	Std. Err.	ANOVA		SD Test		N
					F	Prob > F	f	Pr (F>f)	
<i>Benchmark</i>	No	67.06	12.22	0.92	27.22	0.0000	0.92	0.6050	176
	Yes	73.98	12.71	0.95					178
<i>Justify</i>	No	69.14	13.30	0.99	4.31	0.0387	1.15	0.3481	180
	Yes	71.98	12.39	0.94					174
<i>By Treatment Group (focus Benchmark)</i>									
Group	Mean	Std. Dev.	Std. Err.	ANOVA		SD Test		N	
				F	Prob > F	f	Pr (F>f)		
1	66.92	12.81	1.37	4.92	0.0279	0.90	0.6322	88	
3	71.27	13.48	1.41					92	
2	67.20	11.67	1.24	31.07	0.0000	1.09	0.6981	88	
4	76.87	11.19	1.21					86	
<i>By Treatment Group (focus Justify)</i>									
Group	Mean	Std. Dev.	Std. Err.	ANOVA		SD Test		N	
				F	Prob > F	f	Pr (F>f)		
1	66.92	12.81	1.37	0.02	0.8780	1.20	0.3865	88	
2	67.20	11.67	1.24					88	
3	71.27	13.48	1.41	9.02	0.0031	1.45	0.0835	92	
4	76.87	11.19	1.21					86	
<i>Control vs Both Treatments</i>									
Group	Mean	Std. Dev.	Std. Err.	ANOVA		SD Test		N	
				F	Prob > F	f	Pr (F>f)		
1	66.92	12.81	1.37	29.72	0.0000	1.31	0.2124	88	
4	76.87	11.19	1.21					86	

Moving on to an analysis of how respondents processed specific performance metrics, three separate one-way ANOVA tests suggest a statistically significant difference between the observed satisfaction variable and the performance assessment respondents provided; in the full sample ($F(10(343)) = 6.49, p = 0.0000$), in the sub-sample that did not see a performance benchmark ($F(10(165)) = 3.62, p = 0.0000$), and in the benchmark sub-sample ($F(10(167)) = 4.19, p = 0.0000$). Knowing this, I can dig a little deeper into how respondents processed different performance metrics.

Table 8 shows the substantive differences in the means of three different sub-groups based upon the performance metric their respondents saw in relation to the performance benchmark (note, this only pertains to Group 3 and Group 4). These data appear to show that the justification prompt led respondents in Group 4 to provide higher performance assessments compared to Group 3 (this was contrary to the initial expectation that a justification requirement would induce System 2 thinking and thus reduce cognitive bias [i.e., the role of the performance benchmark as a reference point]). For example, respondents in Group 4 provided a mean performance assessment more than 5 points larger than those who saw the same information in Group 3 but were not asked to make a justification of the performance assessment.

Table 8 - Mean Performance Assessment Relative to Performance Benchmark									
Relation to Benchmark	<i>Full Sample (Groups 3 and 4)</i>			<i>Group 3</i>			<i>Group 4</i>		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Less than	68.85	12	82	67.96	13.56	46	70	9.72	36
Equal to	71.00	14.06	16	65.1	11.59	10	80.83	12.86	6
Greater than	79.83	10.63	80	77.22	11.83	36	81.95	9.12	44
Total	73.98	12.71	178	71.27	13.48	92	76.87	11.19	86

Table 9 shows results for group means and standard deviations for the full sample and two sub-samples: those who did not see the performance benchmark and those who did. This provides more granularity in my comparisons of the role of the performance benchmark. I will discuss a few points that stand out from this table before moving on to another discussion of statistical significance. First, broadly, I have even more detailed evidence that respondents assessed performance differently if they saw a performance benchmark. For example, if I look at those who saw an observed value of 67, those who saw this in the context of the performance benchmark reported a mean performance of more than five points higher than the group that did not see a performance benchmark. Additionally, in the benchmark sub-sample, a performance metric of 67 appears to be an inflection point. But, in the sub-sample that did not see the benchmark the first observed value that produced a statistically different interpretation from other values within this

sub-sample was ‘70’. This may actually suggest respondents were responding to another cognitive bias that can influence the processing of performance information, a left-most digit bias (Olsen, 2013). Additionally, the fact that the first significantly different observed value was higher in the performance benchmark again suggests the benchmark inflated the interpretation of performance information.

Table 9 - Organizational Performance by Observed Satisfaction									
<i>Observed Variable (Embedded Data)</i>	<i>Performance Assessment (Full Sample)</i>			<i>Performance Assessment (No Benchmark Sub-sample)</i>			<i>Performance Assessment (Benchmark Sub-sample)</i>		
Value	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
62	65.60	9.66	35	61.67	7.70	18	69.76	9.97	17
63	65.66	11.26	35	64.70	11.43	20	66.93	11.30	15
64	65.50	12.89	30	59.67	15.27	12	69.39	9.62	18
65	65.60	13.65	40	61.35	11.80	20	69.85	14.32	20
66	67.96	12.96	23	68.45	9.89	11	67.50	15.70	12
67	<i>68.18</i>	<i>14.84</i>	<i>34</i>	<i>65.67</i>	<i>15.45</i>	<i>18</i>	<i>71.00</i>	<i>14.06</i>	<i>16</i>
68	72.76	12.04	37	66.29	10.59	17	78.25	10.54	20
69	72.33	11.77	24	68.54	13.10	13	76.82	8.47	11
70	78.24	10.88	29	76.18	11.98	11	79.50	10.30	18
71	79.43	10.26	28	74.07	6.86	15	85.62	10.22	13
72	76.56	10.88	39	74.00	9.36	21	79.56	12.00	18
Total	70.54	12.92	354	67.06	12.22	176	73.98	12.71	178
Mean performance assessment for respondents who saw the respective business satisfaction value. Table includes means for the whole sample and sub-samples split by whether respondents saw a performance benchmark. Performance benchmark (67) variable shown in italics. For each sub-sample, the first between-group statistical difference is shown in bold.									

DISCUSSION

The experiments in this study are meant to open our understanding of where it is in the cognitive process of performance information use that cognitive bias may influence the way that public managers use performance information. To my knowledge, this questions has remained unexplored in the performance management literature.

Previous scholarship on performance information use by public managers can be characterized into two camps. A majority of this work adopts an organizational theory perspective

and is highlighted by Moynihan's Interactive Dialogue Model, one of the leading models that scholars use to think about how public managers use performance information. Another growing body of scholarship highlights the role of psychology in explaining how public managers use performance information. Unfortunately, to date, little scholarship has attempted to speak across these two bodies of work.

Following a framework of political information use put forward by Gaines and colleagues (2007), I offer the Performance Information Processing Framework (PIPF). In this framework, the acts of being "aware" of performance information—that is, accurately knowing what the information is or says—and interpreting the performance information, represent distinct steps in which cognitive biases may influence the way that individuals (i.e., public managers) process and eventually use performance information. Following this framework, cognitive processing might influence the way that public managers use performance information. Accordingly, my expectation was that the act of interpreting performance information might be prone to the influence of cognitive bias.

I found considerable evidence that respondents were equally "aware" of the information they viewed. While there was variation in the ability of respondents to accurately recall the exact performance metric they saw across the three experiments, I found little evidence of any statistically significant relationship between a respondent's treatment condition and their response when requested to recall the performance information in question. This suggests that *framing manipulations that were intended to elicit various cognitive biases have little influence on a respondent's ability to know what the information is*. Further, any evidence for variation in the interpretations of the performance information observed in these experiments across experimental frames should point to the process of interpretation as being the step in the cognitive processing

of performance information in which individuals are most likely to exhibit cognitive biases. These results provide experimental evidence in support of part of this framework. This can help management scholars explain how and why public managers use performance information.

The PIPF and these findings suggest the importance of incorporating behavioral explanations into our frameworks and models of how we conceptualize how individuals interact with performance information. Also, this work acknowledges the potential for other and future studies to explore what it means to “use” performance information. To expand on these points, a behavioral perspective potentially suggests that performance information use means something different than has been previously considered in the literature. One insight this perspective might provide to public management scholars, is that researchers would need to be more deliberate in how they theorize about the act of performance information use. Is there a direct link between interpretation and the use stage, as the PIPF suggests? Does this depend on the type—i.e., persuasion or making financial decisions for one’s organization—of use?

Limitations

In addition to the contributions of this study, there are some limitations which merit further discussion. To begin, as this is experimental work, the usual caveats pertaining to the generalizability of findings to what happens in practice apply. Though, if we accept the PIPF and the approach undertaken herein, it is also worth mentioning how future scholarship might be able to extend some of the theoretical limitations of this body of scholarship. What might these be?

First, how well does the PIPF capture moving from interpretation to use? It could be that there are other cognitive stages that might influence how public managers use performance

information. This should be explored in future research. Second, how might incorporating organizational theory influence the PIPF? In addition to the previous point, it could very well be that institutional factors in the organizational or political environment (including and in addition to one's organizational role) could influence how individuals process information under certain conditions. Relatedly, it might be that organizational factors influence what steps in the PIPF are relevant to the circumstances of a public manager's decision making process beyond an experimental setting. This logic is already deeply rooted in the IDM.

Third, and relatedly, how well can the framework be incorporated into the Interactive Dialogue Model? The primary critique of the IDM this study raised was that despite the contributions and value of the IDM, there remain limitations to the model. I believe the experimental findings presented here will be of interest to scholars working in this area because of how the findings point to the value of looking for behavioral mechanisms of our empirical findings. Yet, how well does the PIPF fit into the IDM? Or, should scholars consider it as a standalone framework with four distinct models that potentially describe performance information use under different conditions? I personally feel the IDM is robust enough to accommodate the few changes in assumptions necessary to integrate the PIPF into the larger IDM framework. To this point, I could have offered up the PIPF as a distinct perspective. While this may have led to some amount of scholarship interested in this perspective, I felt it was antithetical to the larger intellectual project of trying to understand how public managers *use* performance information. In this sense, I think trying to incorporate the PIPF into the IDM offers more to scholars over the long run.

CONCLUSION

In recent years, public management scholars have given significant attention to the question of how individuals, including public managers, use performance information. Despite some theoretical markers for how to think about this area of research, more effort has been given to describing the subject through empirical work. For this reason, there is a need for public management scholars to seek to develop our understanding of how individuals use performance information in a way that combines theoretical and empirical work from the past two decades. This essay seeks to do this by combining the Interactive Dialogue Model as a framework for thinking about how public managers use performance information and the recent empirical work that demonstrates the behavioral foundations of performance information use. I present a framework—the performance information processing framework (PIPF)—that seeks to describe the cognitive process of performance information use over four distinct cognitive processing models. I believe this framework can be fully incorporated into the Interactive Dialogue Model with only slight modifications to the assumptions of the IDM.

A growing body of empirical evidence suggests cognitive biases can play a significant role in how individuals use performance information. Yet, heretofore, we have had a limited understanding of a) why that is, b) when it occurs, or c) how we might be able to design performance management systems in a way that mitigates these biases. The PIPF and the empirical results found in this essay help to uncover some of those mysteries. Specifically, I show that individuals, at least in an experimental setting, show similarities in understanding what the performance information they see says. Thus, any evidence for differences in assessments of performance information will likely come from the process of interpreting the raw performance information. That is, cognitive bias is most likely to influence individuals not in their ability to know what performance information is, but in their ability to interpret what it says.

Finally, this research should speak to students of behavioral public administration in several ways. First, it should highlight the need and value in looking at behavioral mechanisms as framework for thinking about how to incorporate psychological insights into public management scholarship. Second, it suggests that theorizing about behavioral constructs and processes may offer a bevy of opportunities to better understand our phenomena of interest. Third, when it comes to public management, more effort is needed in understanding how the interplay of behavioral and organizational theories matters for managing in the public sector.

References

- Ammons, D. N., & Rivenbark, W. C. (2008). Factors influencing the use of performance data to improve municipal services: Evidence from the North Carolina benchmarking project. *Public Administration Review*, 68(2), 304-318.
- Andersen, S. C., & Hjortskov, M. (2015). Cognitive Biases in Performance Evaluations. *Journal of Public Administration Research and Theory*, 26(4), 647-662. doi:doi:10.1093/jopart/muv036
- Andersen, S. C., & Moynihan, D. P. (2016). How Leaders Respond to Diversity: The Moderating Role of Organizational Culture on Performance Information Use. *Journal of Public Administration Research and Theory*, 1-13. doi:doi:10.1093/jopart/muv038
- Anderson, D. M., & Edwards, B. C. (2015). Unfulfilled Promise: Laboratory experiments in public management research. *Public Management Review*, 17(10), 1518-1542.
- Askim, J., Johnsen, Å., & Christophersen, K.-A. (2007). Factors behind organizational learning from benchmarking: Experiences from Norwegian municipal benchmarking networks. *Journal of Public Administration Research and Theory*, 18(2), 297-320.
- Baekgaard, M., Christensen, J., Dahmann, C. M., Mathiasen, A., & Petersen, N. B. G. (2017). The role of evidence in politics: Motivated reasoning and persuasion among politicians. *British Journal of Political Science*, 1-24.
- Baekgaard, M., & Serritzlew, S. (2016). Interpreting performance information: Motivated reasoning or unbiased comprehension. *Public Administration Review*, 76(1), 73-82.
- Barrows, S., Henderson, M., Peterson, P. E., & West, M. R. (2016). Relative Performance Information and Perceptions of Public Service Quality: Evidence From American School Districts. *Journal of Public Administration Research and Theory*, muw028.
- Bartels, L. M. (2002). Beyond the running tally: Partisan bias in political perceptions. *Political Behavior*, 24(2), 117-150.
- Behn, R. D. (1995). The big questions of public management. *Public Administration Review*, 313-324.
- Behn, R. D. (2003). Why measure performance? Different purposes require different measures. *Public Administration Review*, 63(5), 586-606.
- Berman, E., & Wang, X. (2000). Performance measurement in US counties: Capacity for reform. *Public Administration Review*, 60(5), 409-420.
- Bertelli, A. M., & Lynn, L. E. (2006). *Madison's Managers: Public administration and the Constitution*. Baltimore, MD: Johns Hopkins University Press.
- Bouckaert, G., & Van Dooren, W. (2009). Performance Measurement and Management in Public Sector Organizations. In T. Bovaird & E. Löffler (Eds.), *Public Management and Governance* (Vol. 2, pp. 151-164). New York: Routledge.
- Bourdeaux, C., & Chikoto, G. (2008). Legislative influences on performance management reform. *Public Administration Review*, 68(2), 253-265.
- Bouwman, R., & Grimmelikhuijsen, S. (2016). Experimental public administration from 1992 to 2014: A systematic literature review and ways forward. *International Journal of Public Sector Management*, 29(2), 110-131.
- Christensen, J., Dahmann, C. M., Mathiasen, A. H., Moynihan, D. P., & Petersen, N. B. G. (2018). How do elected officials evaluate performance? Goal preferences, governance preferences, and the process of goal reprioritization. *Journal of Public Administration Research and Theory*, 28(2), 197-211.
- Cox, C. (1999). *Nietzsche: Naturalism and Interpretation*.

- Epley, N., & Gilovich, T. (2016). The mechanics of motivated reasoning. *Journal of Economic Perspectives*, 30(3), 133-140.
- Gaines, B. J., Kuklinski, J. H., Quirk, P. J., Peyton, B., & Verkuilen, J. (2007). Same facts, different interpretations: Partisan motivation and opinion on Iraq. *Journal of Politics*, 69(4), 957-974.
- George, B., Desmidt, S., Nielsen, P., & Baekgaard, M. (2016). *Rational planning and politicians' attitudes to spending and reform: replication and extension of a survey experiment*. Paper presented at the 2016 IRSPM Annual Conference.
- Gerber, A., & Green, D. (1999). Misperceptions about perceptual bias. *Annual Review of Political Science*, 2(1), 189-210.
- Gerber, A., & Green, D. P. (1998). Rational learning and partisan attitudes. *American Journal of Political Science*, 794-818.
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology*, 62, 451-482.
- Giles, J. (2011). Social science lines up its biggest challenges. *Nature*, 470, 18-19.
- Grimmelikhuijsen, S., Jilke, S., Leth Olsen, A., & Tummers, L. (2016). Behavioral Public Administration: Combining Insights from Public Administration and Psychology. *Public Administration Review*. doi:10.1111/puar.12609
- Hammerschmid, G., Van de Walle, S., & Stimac, V. (2013). Internal and external use of performance information in public organizations: results from an international survey. *Public Money & Management*, 33(4), 261-268.
- Harvey, D. (2001). *Spaces of Capital: Towards a Critical Geography*. New York: Routledge.
- Ho, A. T.-K. (2006). Accounting for the value of performance measurement from the perspective of Midwestern mayors. *Journal of Public Administration Research and Theory*, 16(2), 217-237.
- Hochschild, J. L., & Einstein, K. L. (2015). Do facts matter? Information and misinformation in American politics. *Political Science Quarterly*, 130(4), 585-624.
- Hvidman, U., & Andersen, S. C. (2016). Perceptions of Public and Private Performance: Evidence from a Survey Experiment. *Public Administration Review*, 76(1), 111-120.
- James, O., Jilke, S. R., & Ryzin, G. G. V. (Eds.). (2017). *Experiments in Public Management Research: Challenges and Contributions*. New York: Cambridge University Press.
- James, O., & Van Ryzin, G. G. (2016). 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.
- Jilke, S., Van de Walle, S., & Kim, S. (2016). Generating usable knowledge through an experimental approach to public administration. *Public Administration Review*, 76(1), 69-72.
- Julnes, P. d. L., & Holzer, M. (2001). Promoting the utilization of performance measures in public organizations: An empirical study of factors affecting adoption and implementation. *Public Administration Review*, 61(6), 693-708.
- Kahneman, D. (2002). Maps of bounded rationality: A perspective on intuitive judgment and choice. In *Nobel prize lecture* (Vol. 8, pp. 449-489).
- Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
- Kroll, A. (2015a). Drivers of Performance Information Use: Systematic Literature Review and Directions for Future Research. *Public Performance & Management Review*, 38(3), 459-486. doi:10.1080/15309576.2015.1006469

- Kroll, A. (2015b). Explaining the Use of Performance Information by Public Managers: A planned-behavior approach. *The American Review of Public Administration*, 201-215.
- Kunda, Z. (1990). The Case for Motivated Reasoning. *Psychological Bulletin*, 108(3), 480-498.
- Litman, L., Robinson, J., & Abberbock, T. (2016). TurkPrime. com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior research methods*, 1-10.
- Löffler, E. (2001). Quality awards as a public sector benchmarking concept in OECD member countries: some guidelines for quality award organizers. *Public Administration and Development*, 21(1), 27-40.
- Meier, K. J., & Bohte, J. (2007). *Politics and the Bureaucracy: Policymaking in the Fourth Branch of Government* (5th ed.). Belmont, CA: Thomson Wadsworth.
- Meier, K. J., Favero, N., & Zhu, L. (2015). Performance gaps and managerial decisions: A Bayesian decision theory of managerial action. *Journal of Public Administration Research and Theory*, 25(4), 1221-1246.
- Moynihan, D. P. (2005). Goal-based learning and the future of performance management. *Public Administration Review*, 65(2), 203-216.
- Moynihan, D. P. (2008). *The Dynamics of Performance Management: Constructing Information and Reform*. Washington, DC: Georgetown University Press.
- Moynihan, D. P. (2010). The Promises and Paradoxes of Performance-Based Bureaucracy. In R. F. Durant (Ed.), *The Oxford Handbook of American Bureaucracy*. New York: Oxford University Press.
- Moynihan, D. P. (2015). Uncovering the Circumstances of Performance Information Use Findings from an Experiment. *Public Performance & Management Review*, 39(1), 33-57.
- Moynihan, D. P., & Hawes, D. P. (2012). Responsiveness to reform values: The influence of the environment on performance information use. *Public Administration Review*, 72(s1), S95-S105.
- Moynihan, D. P., & Ingraham, P. W. (2004). Integrative Leadership in the Public Sector A Model of Performance-Information Use. *Administration & Society*, 36(4), 427-453.
- Moynihan, D. P., & Landuyt, N. (2009). How do public organizations learn? Bridging cultural and structural perspectives. *Public Administration Review*, 69(6), 1097-1105.
- Moynihan, D. P., & Lavertu, S. (2012). Does involvement in performance management routines encourage performance information use? Evaluating GPRA and PART. *Public Administration Review*, 72(4), 592-602.
- Moynihan, D. P., Nielsen, P. A., & Kroll, A. (2017). Managerial Use of Performance Data by Bureaucrats and Politicians. In O. James, S. R. Jilke, & G. G. V. Ryzin (Eds.), *Experiments in Public Management Research: Challenges and Contributions* (pp. 244-269). New York: Cambridge University Press.
- Moynihan, D. P., Nielsen, P. A., & Kroll, A. (2017). Managerial Use of Performance Data by Bureaucrats and Politicians. In O. James, S. R. Jilke, & G. G. V. Ryzin (Eds.), *Experiments in Public Management Research: Challenges and Contributions* (pp. 244-269). New York: Cambridge University Press.
- Moynihan, D. P., & Pandey, S. K. (2010). The big question for performance management: Why do managers use performance information? *Journal of Public Administration Research and Theory*, muq004.
- Moynihan, D. P., & Pandey, S. K. (2010). The Big Question for Performance Management: Why Do Managers Use Performance Information? *Journal of Public Administration Research and Theory*, 20(4), 849-866. doi:10.1093/jopart/muq004

- Moynihan, D. P., Pandey, S. K., & Wright, B. E. (2012a). Prosocial values and performance management theory: Linking perceived social impact and performance information use. *Governance*, 25(3), 463-483.
- Moynihan, D. P., Pandey, S. K., & Wright, B. E. (2012b). Setting the table: How transformational leadership fosters performance information use. *Journal of Public Administration Research and Theory*, 22(1), 143-164.
- Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, 110(3), 472-489.
- Nielsen, P. A. (2013). *Performance Information in Politics and Public Management: Impacts on Decision Making and Performance*. (PhD). Aarhus University, Aarhus, Denmark.
- Nielsen, P. A. (2013). Performance Management, Managerial Authority, and Public Service Performance. *Journal of Public Administration Research and Theory*, 24(2), 431-458. doi:doi: 10.1093/jopart/mut025
- Nielsen, P. A. (2014). Learning from Performance Feedback: Performance Information, Aspiration Levels, and Managerial Priorities. *Public Administration*, 92(1), 142-160. doi:10.1111/padm.12050
- 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. doi:10.1093/jopart/mut051
- Nielsen, P. A., & Moynihan, D. P. (2016). How Do Politicians Attribute Bureaucratic Responsibility for Performance? Negativity Bias and Interest Group Advocacy. *Journal of Public Administration Research and Theory*.
- Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2), 303-330.
- Olsen, A. L. (2013). Leftmost-digit-bias in an enumerated public sector? An experiment on citizens' judgment of performance information. *Judgment and Decision Making*, 8(3), 365.
- Olsen, A. L. (2014). *The Appropriate Response to Performance Information by Citizens: A Large-scale Experiment with Local Politicians*.
- Olsen, A. L. (2015a). Citizen (dis) satisfaction: An experimental equivalence framing study. *Public Administration Review*, 75(3), 469-478.
- Olsen, A. L. (2015b). The Numerical Psychology of Performance Information: Implications for Citizens, Managers, and Policymakers. *Public Performance & Management Review*, 39(1), 100-115.
- Olsen, A. L. (2017). Compared to What? How Social and Historical Reference Points Affect Citizens' Performance Evaluations. *Journal of Public Administration Research and Theory*. doi:<https://doi.org/10.1093/jopart/mux023>
- Oppenheimer, D. M., & Kelso, E. (2015). Information processing as a paradigm for decision making. *Annual review of psychology*, 66, 277-294.
- Priem, R. L., Walters, B. A., & Li, S. (2011). Decisions, decisions! How judgment policy studies can integrate macro and micro domains in management research. *Journal of Management*, 37(2), 553-580.
- Radin, B. (2006). Challenging the performance movement. *Washington DC, Georgetown*.
- Salge, T. O. (2011). A behavioral model of innovative search: Evidence from public hospital services. *Journal of Public Administration Research and Theory*, 21(1), 181-210.
- Schaffner, B. F., & Roche, C. (2016). Misinformation and motivated reasoning: Responses to economic news in a politicized environment. *Public opinion quarterly*, 81(1), 86-110.

- Seidenfeld, M. (1996). Demystifying deossification: Rethinking recent proposals to modify judicial review of notice and comment rulemaking. *Tex. L. Rev.*, 75, 483.
- Seidenfeld, M. (2001). The Psychology of Accountability and Political Review of Agency Rules. *Duke Law Journal*, 51(3), 1059-1095. doi:10.2307/1373184
- Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, 50(3), 755-769.
- Taylor, J. (2009). Strengthening the link between performance measurement and decision making. *Public Administration*, 87(4), 853-871.
- Van de Walle, S., & Van Dooren, W. (2011). Introduction: Using Public Sector Performance Information. In W. Van Dooren & S. Van de Walle (Eds.), *Performance Information in the Public Sector: How it is Used* (pp. 1-12). New York: Palgrave Macmillan.
- Van Dooren, W., & Van de Walle, S. (Eds.). (2011). *Performance Information in the Public Sector: How it is Used*. New York: Palgrave Macmillan.
- West, W. F. (1995). *Controlling the Bureaucracy: The Theory and Practice of Institutional Constraints*. Armonk, NY: M.E. Sharpe.
- West, W. F. (2004). Formal procedures, informal processes, accountability, and responsiveness in bureaucratic policy making: An institutional policy analysis. *Public Administration Review*, 64(1), 66-80.
- Yang, K., & Hsieh, J. Y. (2007). Managerial effectiveness of government performance measurement: testing a middle-range model. *Public Administration Review*, 67(5), 861-879.
- Yeo, S. K., Cacciatore, M. A., & Scheufele, D. A. (2015). News selectivity and beyond: Motivated reasoning in a changing media environment. In *Publizistik und gesellschaftliche Verantwortung* (pp. 83-104): Springer.
- Yockey, M. D., & Kruml, S. M. (2009). Everything is relative, but relative to what? Defining and identifying reference points. *Journal of Business and Management*, 15(1), 95.

Appendix

Experimental Vignettes and Workflow			
Stages	No Benchmark		Benchmark
Introduction	For this question, imagine that you are the manager of a business development office for a major metropolitan area. Your city just released its yearly performance metrics and, based on this information, the mayor wants to know how you think the city performed over the course of the last year.		
Question Frame (Benchmark randomization)	For some time, business owners were asked whether they were generally satisfied or generally dissatisfied with the city as a place to do business. In 2011, 57% of business owners indicated they were satisfied with the city as a place to do business. The mayor has tasked your office with improving the business climate in the city.		For some time, business owners were asked whether they were generally satisfied or generally dissatisfied with the city as a place to do business. In 2011, 57% of business owners indicated they were satisfied with the city as a place to do business. The mayor has tasked your office with improving the business climate in the city. The goal has been to increase the percentage of business owners satisfied with doing business in the city to 67%.
Justify	No Justify	In the following question you will be given performance information. You will then be asked to provide an assessment of the city's performance over the last year given this performance information. Please only consider the information before you at that time when providing an assessment. After you provide the assessment, on the following screen you will be asked to provide a justification for the performance assessment you just provided.	No Justify
Performance Metric Question (XX% signifies a randomly generated performance metric.)	XX% of business owners were satisfied with the city as a place to do business. Using the information available to you, use the sliding scale (0-100) to rate the city's performance in regard to business development based upon this performance information:		XX% of business owners were satisfied with the city as a place to do business. Remember, the mayor's goal was that 67% of business owners would be satisfied with the city as a place to do business. Using the information available to you, use the sliding scale (0-100) to rate the city's performance in regard to business development based upon this performance information:

Randomization Checks		
Variable	F	Prob > F
Gender	0.01	0.9987
PSM	0.90	0.4423
PSM_APM	1.17	0.3208
PSM_COM	0.94	0.4205
PSM_CPI	1.52	0.2097
PSM_SS	0.50	0.6848
Experience (Overall)	1.73	0.1596
Experience (Public Sector)	2.09	0.1017
<p>This table provides the results of a randomization check for Experiment 2. For each treatment group, we provide the mean afor each of eight potential control variables. We ran ANOVA tests on each of the controls as a check on randomization. Of note, each of the controls appears to pass this check at the $p = 0.05$ level.</p>		
<p>Variables: PSM (Public service motivation), PSM_APM (Attraction to policy making), PSM_CPI (Commitment to public interest), PSM_COM (Compassion), PSM_SS (Self-sacrifice); Kim (2011) and Perry (1996).</p>		