

UNIVERSITA' COMMERCIALE "LUIGI BOCCONI"

PhD SCHOOL

PhD program in Public Policy and Administration

Cycle: 31

Disciplinary field: SPS/04

**Systematic deviations from rational
decision-making in public administration**

Advisor: Anthony BERTELLI

Co-Advisor: Nicola BELLÉ

PhD Thesis by
Paolo BELARDINELLI
ID number: 1610476

Academic Year 2019/2020

Acknowledgements

Many people contributed to make my research path far easier than I expected when I started to think about applying to a PhD program: all my teachers, all my coauthors, all the colleagues in my PhD cohort, and many others inside and outside Bocconi University. I thank them all very much, indeed. If I were to express adequately my gratitude towards all of them, the names I would need to list and the words I would need to spend should be so many that I would look ridiculous. This would be a terrible start of a PhD dissertation.

For this reason, I shall confine myself to explicitly thanking current and past members of my committee here at Bocconi: Maria Cucciniello as a current member, Ross Macmillan as a past member and former PhD director, and Valentina Mele as a permanent member who has been supportive throughout the entire process and with whom I was engaged in multiple stimulating collaborations. I also thank Al Rainey and Gregg Van Ryzin for their inspiring comments in the reviewing process that will keep me busy for the next few years. Among the others, I am very grateful to Giovanni Fattore, who, besides useful suggestions, provided me with resources and freedom at a moment when I needed them both.

Tony Bertelli has been a great advisor since the very first moment he arrived at Bocconi; when I kind of lost the perception of a clear path, he became a new reference point. Nicola Bellé, who appears here as external co-advisor, contributed more than anyone else to enrich my research path, since the very beginning when I was still a master student. I hope that I will never give him the opportunity to regret all he taught me.

Introduction

When making judgments and decisions, individuals have been widely shown to be prone to cognitive biases, leading them to inconsistently reveal their preferences (e.g. Kahneman 2011). Nonetheless, despite the fact that these can have far-reaching consequences for our society, research on how systematic errors impinge on the decisions made within the public sector is limited.

The three chapters included in this proposal provide empirical evidence, by means of randomized experiments, aimed at increasing our knowledge of these dynamics. In particular, the first chapter replicates and extends previous trials by testing a broad range of cognitive biases on public policy and management decisions on a sample of public managers and employees. While this chapter is purposefully wide-ranging, the remaining two dig deeper on frequent decision-making processes of public workers. More specifically, the second chapter focuses on performance information use by Italian local public managers, showing that they are more likely to be subject to framing effects under *ex post* uses of performance information (e.g., service evaluation) than *ex ante* (e.g., resource allocation). The third chapter investigates the micro-foundations of isomorphic pressures (DiMaggio and Powell 1983) which lead to suboptimal decision-making in the public sphere, adding qualitative evidence which helps illuminate the mechanisms behind the causal relations.

“To an economist *rationality* is consistency and nothing more” (Caplin and Glimcher 2014, p. 11). Being cognitive biases broadly defined as systematic deviations from whatever normative idea of rationality (Larrick 2004), here cognitive biases are more specifically considered to be systematic deviations from consistency in revealed preferences. In other words, the focus of this work is on violations of what Kahneman (2000) defines as coherence rationality, which is “the ability to reason correctly about immediately available information” (Kahneman 2000, 682).

Understanding how supposedly irrelevant factors of the choice architecture may alter public decision-making in predictable ways (Thaler and Sustein 2008; Thaler 2017) is increasingly catalyzing scholarly attention (Kelman, Sanders, and Pandit 2016; Vlaev et al. 2016; Moynihan, Herd, and Harvey 2014; Gordon, Kornberger, and Klegg 2009). A robust discussion has begun on decision-making biases in public management, administration, and policy (e.g., among the most recent Andersen and Hjortskov 2015; Baekgaard and Serritzlew 2016; Geys and Sørensen 2017; Grosso, Charbonneau, and Van Ryzin 2017; Jilke, Van Ryzin, and Van de Walle 2016; Olsen 2017). These scholars have mostly focused on how citizens make informed assessments of government policies while only few of them have investigated how public managers and policy makers make decisions (e.g., Meier, Winter, O’Toole, Favero, and Andersen 2015; Bellé, Cantarelli, and Belardinelli 2017).

Nonetheless, it is equally relevant and timely to understand whether and how cognitive biases affect public workers responsible for designing, managing, and implementing public policies and managerial procedures. John and Stoker (2019) recently argued that, in trying to design effective policies, “there are reasons to doubt the credentials of experts, while not denying them a role” (p. 216). The work included in the three chapters shows the importance of taking deviations from rational decision making among public managers and employees into account while thinking about policy interventions and management procedures. This seems imperative also in light of Herbert Simon’s early argument that one of the main functions of government organizations is to cope with the limits of their members’ “abilities to comprehend and compute in the face of complexity and uncertainty” (Simon 1978, p. 345).

1 Prospect Theory Goes Public: Experimental Evidence on Cognitive Biases in Public Policy and Management Decisions¹

Behavioral sciences suggest that public servants' judgments may be systematically biased under certain circumstances. This conflicts with decision-making models that are based on expected utility theory, which features rational agents (Bernoulli 1954). In this article, we investigate how public managers and employees actually make decisions by putting to an experimental test a number of cognitive biases that branch out from or are associated with prospect theory (Kahneman 2011; Kahneman and Tversky 1979). From an epistemological standpoint, our work aims to experimentally test descriptive models "that accurately portray human behavior" (Thaler 2015, 30) in the context of public management and policy. We do so by replicating and extending previous randomized controlled trials.

Whereas behavioral public administration (e.g., Grimmelikhuijsen et al. 2017) studies tend to focus on one or a few cognitive limitations at a time and rely on samples of citizens (e.g., Baekgaard 2017; Baekgaard and Serritzlew 2016; Barrows et al. 2016; Jilke, Van Ryzin, and Van de Walle 2016; Marvel 2016; Olsen 2017), this article tests the effects of a broad range of cognitive biases across multiple policy areas and managerial tasks. To this end, we conducted 10 randomized controlled trials to explore how framing, anchoring, proportion dominance, status quo, and asymmetric dominance affected the decisions of 600 Italian public workers from different industries. We selected these cognitive biases starting from a more comprehensive list and narrowing it down to the ones that we believe have more direct implications for public policy making and public management.

¹This paper was conducted in collaboration with Nicola Bellé and Paola Cantarelli. It has been published in the *Public Administration Review*.

1.1 Theoretical Background

Before the 1940s, expected utility theory was the dominant model used to describe decision making. At its core, expected utility theory features a rational decision maker who has clear and comprehensive knowledge of the environment, a well-organized system of preferences, and excellent computational skills to allow for the selection of optimal solutions (Bernoulli 1954). Starting with the work of Herbert Simon in the 1940s and 1950s, scholars have contended instead that decision makers are endowed with bounded rationality. Simon (1947, 1956) claimed that decision makers do not optimize but rather “satisfice.” Literally a combination of “satisfy” and “suffice,” the satisficing principle suggests that individuals forgo “optimum solutions for a simplified world [in favor of] satisfactory solutions for a more realistic world” (Simon 1978). More precisely, people make decisions for themselves and for others by relying on a limited number of heuristic principles that reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations.

Consistent with the paradigm of bounded rationality, there is abundant evidence that decision makers are prone to cognitive biases that systematically affect their estimates, judgments, and choices in any domain. For example, “the orthodox economic model of consumer behavior is, in essence, a model of robot-like experts. As such, it does a poor job of predicting the behavior of the average consumer” (Thaler 1980, 58). Most of this literature stems from the work of Daniel Kahneman, who distinguishes between system 1 and system 2 thinking. System 1 thinking is synonymous with intuition and is fast, rigid, effortless, automatic, and associative. Regardless of whether it originates from skills or from heuristics, intuition comes to mind easily and confidently (Kahneman and Klein 2009). Reasoning, instead, happens in system 2 thinking, which is slow, controlled, effortful, flexible, and rule governed.

Heuristics can be effective from an ecological perspective because they are shortcuts that reduce complexity (Artinger et al. 2015; Gigerenzer and Todd 1999; Smith 2003). Yet heuristics may also lead to systematic errors (e.g., Ariely 2010; Kahneman 2002, 2011). Drawing on Thaler (2015, 2017) and Thaler and Sunstein (2008), we consider a choice to be systematically biased when the judgment made by a decision maker differs from the judgment that a robot would have made. In the words of Nobel Prize laureate Richard H. Thaler, “standard economics assumes that agents in the economy are rational maximizers who are more like robots than humans” (Katz 2017). For instance, when asked to estimate a certain quantity, a robot would not be influenced by a small or large numeric anchor but rather would provide exactly the same estimate. To the contrary, human decision makers tend to give estimates that are systematically biased toward the low numeric anchor or toward the high numeric anchor, depending on which of the two they have been randomly exposed to.

Similarly, robots would not make different decisions based on whether information is provided in a positive rather than a negative manner. Instead, in a classic experiment conducted at the Harvard Medical School by Amos Tversky and colleagues, physicians were 84 percent likely to perform a surgery when told that the one-month survival rate was 90 percent but only 50 percent likely to perform the same surgery when informed that there was a 10 percent mortality in the first month because “the statement that the odds of survival one month after surgery are 90% is more reassuring than the equivalent statement that mortality within one month of surgery is 10%” (Kahneman 2011, 88). In other words, as suggested by an anonymous reviewer, our study investigates violations of what Kahneman (2000) defines as coherence rationality, which is “the ability to reason correctly about immediately available information” (Kahneman 2000, 682), rather than reasoning rationality.

The interest in understanding the effects of cognitive biases is widespread across disciplines. Some studies have focused on one particular cognitive bias at a time, exploring its effect on either managerial (e.g., Bechger, Maris, and Hsiao 2010; Chen and Kemp 2015; Jacobs and Kozlowski 1985; O'Donnell and Schulz 2005) or generic decision tasks (e.g., Ariely 2010; Tversky and Kahneman 1974; Wu and Cheng 2011). Other scholars have investigated how multiple biases influence decisions in specific professions (e.g., Baekgaard et al. 2017; Blumenthal-Barby and Krieger 2015; Butler and Broockman 2011; Englich, Mussweiler, and Strack 2006; Guthrie, Rachlinski, and Wistrich 2001; Linde and Vis 2017; Rachlinski, Guthrie, and Wistrich 2006; Vis 2011). Literature reviews represent another stream of scholarship on cognitive biases (e.g., Bennett 2014; Carter, Kaufmann, and Michel 2007; Cornelissen and Werner 2014; Furnham and Boo 2011; Kühberger 1998; Orr and Guthrie 2006). Most of this scholarship traces back to prospect theory (Kahneman and Tversky 1979) or associated work on system 1 and system 2 modes of thinking (Kahneman 2011).

Unlike in other disciplines such as applied psychology (e.g., Kahneman 2011), general management (e.g., Cornelissen and Werner 2014), marketing (e.g., Ariely 2010), and medicine (Blumenthal-Barby and Krieger 2015; Saposnik et al. 2016), which have a long tradition in the study of cognitive biases, this area of research is still nascent in public administration (Grimmelikhuijsen et al. 2017). Scholars in our field have only recently conducted experimental research into how cognitive biases affect citizens' judgments of public services (e.g., Andersen and Hjortskov 2015; Baekgaard 2017; Baekgaard and Serritzlew 2016; Barrows et al. 2016; Jilke, Van Ryzin, and Van de Walle 2016; Marvel 2015, 2016; Olsen 2015, 2017). Similarly, analyzing 159 case studies from 60 public bodies in 23 states and two international institutions, the Organisation for Economic Co-operation and Development recently reported that attempts are underway to use behavioral

insights to inform policies across policy areas (such as consumer behavior, education, energy, environment, finance, health and safety, labor market, service delivery, taxes and telecommunications) (OECD 2017). Research on how systematic errors impinge on the decisions of public officials and public managers is more limited (e.g., Bellé, Cantarelli, and Belardinelli 2017; Meier et al. 2015; Moynihan and Lavertu 2012). Nonetheless, understanding how cognitive biases affect decisions made within public organizations is of key importance for public administration studies, as these decisions can have far-reaching consequences for citizens. If exploring the systematic errors of citizens' judgments about government services is important, it is equally relevant and timely to understand whether and how cognitive biases affect public workers responsible for designing, managing, and applying public policies and managerial procedures.

Our study aims to provide a twofold contribution to research in this area. First, we follow up on recent calls to engage in replication studies (e.g., Bellé and Cantarelli 2017; George et al. 2016; Levitt and List 2007; Walker, James, and Brewer 2017) by focusing on decision-making of public sector workers rather than citizens. Indeed, “replication sit[s] at the heart of scientific progress” (Walker, James, and Brewer 2017, 1221). Second, instead of focusing on one or a few cognitive biases at a time, our work explores the consequences of a broad range of cognitive limitations that may affect public policy and management decisions. This approach echoes previous studies in other fields that have comprehensively investigated how different types of reasoning errors may bias decisions in specific professions, such as for medical doctors (e.g., Blumenthal-Barby and Krieger 2015) and judges (e.g., Guthrie, Rachlinski, and Wistrich 2001).

The following sections provide an explanation of the cognitive biases that we test empirically in our sample: framing, anchoring, proportion dominance, status quo, and asymmetric dominance. We then describe the research design of each

randomized controlled trial, present results, and discuss how sizable deviations from rational decision making might be taken into account while designing policy interventions, management systems, and procedures.

Framing

“Few theoretical constructs have such widespread traction. . . as the construct of frame or framing” (Cornelissen and Werner 2014, 182). Indeed, extant research syntheses (e.g., Cornelissen and Werner 2014; Kühberger 1998) and recent world-renowned books (e.g., Kahneman 2011) document the policy areas and managerial decisions in which framing effects have been demonstrated extensively and consistently. The framing effect tends to occur in decision-making situations that reproduce typical processes of public policy formulation and choice. Therefore, we are convinced that public policy making, which is at the core of what public administration does, could benefit from a deeper understanding of how and to what extent it may be prone to systematic errors caused by cognitive limitations.

The framing effect seems to unfold consistently in decision-making tasks that replicate the elements of the now very famous Asian disease problem, originally developed by Tversky and Kahneman (1981). In these tasks, individuals are required to select one of two public policies. The expected value of the outcome of the two public policies is the same, but whereas the outcome of one policy is expressed as a sure thing, the outcome of the other policy is expressed in probabilistic terms. Therefore, individuals are asked to choose between a sure thing and a gamble. The between-subjects difference of this decision-making task lies in whether the outcome of the two policies has a positive or a negative frame. In one scenario, the consequence of each of the two policies is framed positively (i.e., lives saved in the Asian disease problem), while in the other scenario, consequences are framed negatively (i.e., lives lost in the Asian disease problem). In this situation, the

framing effect consists in a shift of individuals' risk preferences. More precisely, abundant experimental evidence has shown that individuals prefer the policy with the sure outcome when the outcomes are framed positively and prefer the policy with the probabilistic outcome when outcomes are framed negatively. In other words, under the framing effect, decision makers are risk averse in the domain of gains and risk takers in the domain of losses (Tversky and Kahneman 1981).

The framing effect, then, is a cognitive bias, whereby individuals tend to react in a systematically different manner to the same piece of information, depending on how it is presented to them—for instance, as the percentage of customers satisfied with a service as opposed to the percentage of unsatisfied customers, or as the survival rate rather than the equivalent mortality rate. In other words, objectively equivalent information may lead to systematically different behavioral outcomes, depending on the framing of the information (e.g., Olsen 2015).

To test the external validity of the findings from previous research on Italian public managers and employees in different decision scenarios, we designed five survey experiments in which we manipulated the framing of information. We expected participants in our experiments to behave in a systematically different manner depending on whether they were exposed to positively or negatively framed information.

Anchoring

Anchoring is the cognitive tendency to estimate unknown quantities by making adjustments from an initial value. Experimental studies have consistently shown that “different starting points yield different estimates, which are biased toward the initial values” (Tversky and Kahneman 1974, 1128). In other words, decision makers asked to assess an unknown quantity and consider a certain number for their evaluation tend to provide a final estimate that is an insufficient adjustment

of the initial value (e.g., Kahneman 2011; Tversky and Kahneman 1974).

Our study of the anchoring effect in a core function of public administration tends to align very nicely with extant calls to understand the serious challenges posed by the big data revolution in government (e.g., Lavertu 2016; Moynihan and Pandey 2010). Despite being “the most recent manifestation of continued efforts to base public decision-making on measured quantities and to enhance transparency and accountability” (Lavertu 2016, 865), our knowledge of whether and how these data may serve as anchors for public decision makers is still underdeveloped.

Classic studies have demonstrated that incomplete computations, numbers generated randomly in the presence of the decision maker, and numbers provided with a priming role function as anchors and lead to the anchoring effect (e.g., Furnham and Boo 2011; Wilson et al. 1996). Extant experimental research has extensively and consistently found the anchoring effect in domains as different as general knowledge (e.g., Simmons, LeBoeuf, and Nelson 2010), quantitative evaluation of employees’ performance (e.g., Bellé, Cantarelli, and Belardinelli 2017; Thorsteinson et al. 2008), suggestions for employees’ promotion (e.g., Chen and Kemp 2015); legal judgments (e.g., Bennett 2014), negotiations (e.g., Orr and Guthrie 2006), and economic valuations (e.g., Alevy, Landry, and List 2015). Within public human resource management, Bellé et al. (2017) recently showed how individual performance appraisal is biased toward previous years’ performance scores.

To test the external validity of previous scholarship, we tested whether different anchors led Italian public employees and managers to give different answers about the maximum number of days within which employees in a hypothetical municipality should reply to citizens’ emails. We expected Italian public employees and managers in our sample to indicate a higher (lower) maximum number of days when exposed to a high (low) anchor.

Proportion Dominance

The proportion dominance effect is a cognitive bias whereby individuals tend to put more weight on the percentage rather than the absolute number of people affected by their decisions and actions. Scholars have shown that subjects tend to be more willing to help higher rather than lower percentages of victims or beneficiaries of a service, even when the absolute number is held constant (e.g., Baron 1997; Bartels 2006; Erlandsson, Björklund, and Bäckström 2015; Fetherstonhaugh et al. 1997; Slovic et al. 2007). As an example, the helping motivation of subjects asked to support African children with bacterial meningitis showed a linearly increasing pattern when children numbered 275 out of 8,000 (3 percent), 275 out of 2,000 (14 percent), 275 out of 900 (31 percent), and 275 out of 300 (92 percent), respectively (Erlandsson, Björklund, and Bäckström 2015).

Based on these priors, we expected that public employees in our sample would be willing to dedicate more time to a project affecting a larger percentage of potential beneficiaries compared with a project affecting a smaller percentage of potential beneficiaries, even if the actual number of beneficiaries was the same.

Status Quo

Decision makers tend to prefer the status quo option as the number of viable alternatives increases. The multiple alternative effect has clear implications for public managers, who are routinely faced with choices between sticking to status quo providers or switching to new contractors. To explain this effect, Thomas Schelling told an anecdote about his decision to buy an encyclopedia. At the bookstore, two encyclopedias were on sale. Finding it difficult to choose between the two, Schelling ended up buying neither of them. In the case of a single encyclopedia being on sale, he would have been happy to buy it.

Regardless of what the optimal outcome is, decision makers tend to prefer the

status quo when there are too many attractive alternatives that would improve their condition. Experimental evidence sheds light on what happened to Schelling. Tversky and Shafir (1992), for example, asked students to consider buying a CD player. One group was asked to decide whether to buy a popular CD player, buy a top-of-the-line CD player, or wait until learning more about various models. The other group was presented with the same scenario, without the top-of-the-line CD player option. In the first scenario, 46 percent of the students delayed the purchase, while in the second, 34 percent did. The addition of a competing alternative increased the tendency to delay the decision and prefer the status quo.

Redelmeier and Shafir (1995) tested the same effect on family physicians, who were presented with a scenario involving the choice of whether to start a new medication for a patient. In addition to referring the hypothetical patient to an orthopedic consultant, participants in one group had to choose between starting or not with ibuprofen medication, while participants in the other group had to choose whether to start the ibuprofen medication, start the piroxicam medication, or not start a new medication. In the first case, 53 percent of physicians did not prescribe a new medication. In the second, this share rose to 72 percent. In public administration research, Moynihan and Lavertu (2012) found that the local election officials in their study tended to provide higher ratings for whatever voting technology was in use in their jurisdictions. For instance, “we find that local officials who report having a general faith in technology and use DREs [direct recording electronic] as their main voting system also report a greater preference for DREs compared with other voting systems” (Moynihan and Lavertu 2012, 73).

We designed two experiments to test whether the effect of multiple alternatives held true across different experimental settings. We expected preferences of Italian public managers and employees in our sample to shift toward the status quo as the number of viable alternatives increases.

Asymmetric Dominance

The asymmetric dominance effect, also known as decoy, causes individuals' preferences between options A and B to shift toward option A when a third option A' , similar to A but in no way better, is added to the choice set. The consequence of adding an irrelevant alternative to the choice set is surprising because it generates a shift in preferences that is dependent on the context. Even more surprisingly, the option that causes the systematic rather than random shift in preferences is rarely, if ever, selected.

Evidence of the asymmetric dominance effect has nurtured the development of models of context-dependent preferences (e.g., Huber, Payne, and Puto 1982; Tversky and Simonson 1993). The simplest case refers to decision-making processes in which subjects are required to make a choice among three options (i.e., A , B , or A') that are described along two attributes. Option A is the baseline-targeted option, B is the baseline-competing option, and A' is the decoy option for A . The expression "asymmetric dominance" is due to the fact that A dominates B along one attribute but B is superior to A along the other attribute. Therefore, neither A nor B is univocally dominant.

The decoy option A' is equal to A in one attribute and slightly inferior to A in the other attribute, or it is slightly inferior to A along both attributes. Meanwhile, A' is inferior to B in one attribute but superior to B on the other attribute. These comparative features of the three options make A dominate A' and A' be the decoy for A or, equivalently, the asymmetrically dominated option. The decoy effect predicts that the probability that an individual would prefer A over B is higher when A' is present relative to when A' is absent. In other words, the salience and desirability of one option over another depends on the presence of a third alternative that is suboptimal to one option but not to the other. While the decoy

is virtually never chosen as the best option, it has been shown to influence the choice between the other alternatives in the choice set.

The asymmetric dominance effect has been found in experimental tasks as different as selecting a candidate for a job (e.g., Highhouse 1996; Slaughter, Bagger, and Li 2006; Slaughter, Sinar, and Highhouse 1999), picking a journal subscription option (e.g., Ariely 2010), choosing private or club goods to purchase (e.g., Ariely and Wallsten 1995; Huber, Payne, and Puto 1982; Huber and Puto 1983), and selecting a partner (e.g., Sedikides, Ariely, and Olsen 1999). Most of these are choices that public managers are called on to make with high frequency as part of their jobs.

1.2 Research Design

We administered 10 factorial randomized controlled trials to a sample of 600 Italian public sector workers. Within each trial, subjects were randomly assigned to conditions (i.e., scenarios). Table 1 reports the experimental factors, factor levels, and actual scenarios that subjects read.

Experiment 1 replicated Tversky and Kahneman's (1981) original study of the Asian disease problem. It provides a baseline test of the framing effect by investigating whether the risk aversion of public sector managers and employees varied based on whether sure and probabilistic policy outcomes were framed positively or negatively. We designed experiment 2 to strengthen the external validity of the findings of experiment 1 by varying the policy area. We selected the environmental policy-making domain for experiment 2 following the procedures used by Tversky and Kahneman (1981), who used the Asian disease problem to increase realism by priming respondents' memory about the Asian flu epidemic of 1957. Indeed, not only is Italy regularly hit by major floods, but also a major flood caused damage in Florence a few weeks before our data collection window.

Table 1: Experimental Factors and Scenarios

Experimental factor (levels)	Experiment Number and Scenario
Framing (positive; negative)	<p>1: Imagine that Italy is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:</p> <ul style="list-style-type: none"> • If program A is adopted, <i>200 people will be saved [400 people will die]</i> • If program B is adopted, there is a one-third probability that <i>600 people will be saved [nobody will die]</i> and a two-thirds probability that <i>no people will be saved [600 people will die]</i> <p>Which of the two programs would you favor?</p> <p>2: You are the senior manager of the Civil Protection Department in charge of the national flood prevention program. Experts predict that floods will kill 600 people in the next years. Two alternative programs to prevent floods have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:</p> <ul style="list-style-type: none"> • If program A is adopted, <i>200 people will be saved [400 people will die]</i> • If program B is adopted, there is a one-third probability that <i>600 people will be saved [nobody will die]</i> and a two-thirds probability that <i>no people will be saved [600 people will die]</i> <p>Which of the two programs would you favor?</p> <p>3: You have a manager position at the Department of Health and you need to choose one of the following two programs to combat a disease:</p> <ul style="list-style-type: none"> • If program A is adopted, <i>10,000 people will be saved [20,000 people will die]</i> • If program B is adopted, there is a one-third probability that <i>30,000 people will be saved [nobody will die]</i> and a two-thirds probability that <i>no people will be saved [30,000 people will die]</i> <p>Which of the two programs would you favor?</p> <p>4: You are the principal of a secondary school. You need to decide whether to purchase the electronic school register, Attendance. A survey conducted on a representative sample of students and parents indicates that <i>80% of the respondents are satisfied [20% of the respondents are dissatisfied]</i> with the software, Attendance. How likely are you to purchase Attendance?</p> <p>5: You are the principal of a secondary school. You need to decide whether to purchase the electronic school register, Attendance, and could opt for only satisfied or dissatisfied. <i>80% of the respondents are satisfied [20% of the respondents are dissatisfied]</i>. How likely are you to purchase Attendance?</p>
Anchoring (low; high)	<p>6: You are the senior manager of the Public Relations Office in a medium-sized municipality. You have to decide the maximum number of days by which your subordinates have to reply to citizens' inquiries sent via emails. Consider whether the maximum number of days to reply to citizens' emails must be higher or lower than 2 [90] working days. Indicate the maximum number of days below.</p>
Proportion dominance (low, medium, high)	<p>7: You are the head of Social Services. You have to select the new supplier in charge of delivering hot meals to 275 of the 2,000 [900] [300] senior citizens who are currently provided with the service. In other words, you have to decide the supplier in charge of delivering hot meals to 14% [31%] [92%] of the total beneficiaries of the service. During the next week, what percentage of your working time would you dedicate to the selection of the supplier, taking it away from other equally important tasks?</p>
Status quo (one alternative; two alternatives)	<p>8: Your public employer asks you to consider whether to renew the contract with the current hardware and software contractor or switch. The market analysis that your staff prepared for you identifies <i>Omega Ltd as a viable alternative [Omega Ltd and Sigma Ltd as viable alternatives]</i>. Omega Ltd provides slightly better service as compared to the current contractor at the same cost. <i>[Sigma Ltd provides an equally good service as compared to the current contractor at a slightly lower cost]</i>. Obviously, switching contractors would require time and paperwork. Do you renew the contract with the current contractor?</p> <p>9: You are the Personnel Director in a public organization. You recently noticed problems in the functioning of the payroll software. After several attempts, the current contractor had been unable to fix them. Therefore, you prepare a call for bids to select a new contractor. Immediately before making the call for bids public, you talk to the IT office of your organization and become aware of a <i>software-fix procedure [two-software fix procedures]</i> that could potentially solve the problems. Trying <i>this procedure [these procedures]</i> would require time and there might be delays in the payment of the thirteenth salary (Christmas bonus) if <i>it does [they do]</i> not fix the problems. Do you publish the call for bids without trying the <i>software-fix procedure [procedures]</i>?</p>
Asymmetric dominance (no; yes)	<p>10: Your public employer asks you to choose a subscription plan for the Il Sole 24 Ore. Choose one of the following:</p> <ul style="list-style-type: none"> • Web subscription—59 euros. One-year subscription to <i>ilsole24ore.com</i>. Includes online access to all the articles from Il Sole 24 Ore since 2006 • <i>[Print subscription—125 euros. One-year subscription to the print edition of Il Sole 24 Ore]</i> • Print and web subscription—125 euros. One-year subscription to the print edition of Il Sole 24 Ore and online access to all the articles from Il Sole 24 Ore since 2006

Note: The text in italics displays our experimental manipulations. The text in italics in the square brackets was either inserted instead of the corresponding italics text in the vignette or added to the vignette. For example, in the anchoring experiment, a group of subjects read “Consider whether the maximum number of days to reply to citizens’ emails must be higher or lower than 2 working days” while another random group read “Consider whether the maximum number of days to reply to citizens’ emails must be higher or lower than 90 working days”.

Experiment 3 also addressed potential external validity concerns of the findings of experiment 1. Experiment 1 and experiment 3 shared the same policy area and outcomes but differed in the absolute numbers of expected lives saved and lost. We selected the figures based on the statistics of annual deaths for the most common cancers in Italy.

Experiments 4 and 5 explored the framing effect, building on Olsen (2015). As in Olsen, participants were presented with either a customer satisfaction rate or a customer dissatisfaction rate. Experiments 4 and 5 differed in one element only. Unlike experiment 4, experiment 5 specified that the survey administered to the representative sample of students and parents would feature only two response options: satisfied or dissatisfied. This clarification makes positively and negatively framed information unambiguously equivalent, which is not the case with three or more response options. We designed experiment 4 as our baseline and experiment 5 to rule out the possibility that the findings of experiment 4 were threatened by the potential ambiguity about the number of satisfaction categories.

Experiment 6 tested whether public sector managers and employees were prone to anchoring effects in making decisions about the standards of responsiveness toward citizens. We prompted participants to think about the anchor before providing their response, following the procedures adopted by Tversky and Kahneman (1974). Also, we decided the numeric value of our anchors based on Gilke, Van Ryzin, and Van de Walle (2016), who used the same figures referring to the time that it took a municipality-owned electric company to refund money. We built experiment 7 on the proportion dominance effect study of Erlandsson, Björklund, and Bäckström (2015) and asked public managers and employees to indicate what percentage of their working time they were willing to dedicate to the selection of a new supplier for a public service.

Experiments 8 and 9 tested the multiple alternatives effect. Public managers

and employees in experiment 8 decided whether or not to renew the contract with the current hardware and software contractor. Sticking to the current contractor represented the status quo and served as the dependent variable. More precisely, the dependent variable was a dichotomous variable that took on the value of 1 if respondents decided to stick to the current hardware and software contractor and the value of 0 otherwise. The scenario provided a few details about the alternative contractor(s) to preserve realism. Furthermore, we opted to make the alternative(s) comparable to the current supplier overall to avoid the risk that any observed effect was inflated or driven by the superiority of the alternative(s) rather than by their number. Experiment 9 tested the external validity of the findings of Redelmeier and Shafir (1995), substituting an administrative decision within a public organization for the medical setting.

Experiment 10 aimed to strengthen the external validity of the findings reported in Ariely's 2010 book. More precisely, it replicated the experimental treatments of the *Economist* study using the most important Italian economics daily newspaper (*Il Sole 24 Ore*).

One significant way in which all of our randomized controlled trials departed from the original studies was that subjects were real public sector employees and managers instead of convenient samples of students or citizens. We selected public decision-making scenarios based on two criteria. First, we used vignettes that were as realistic as possible for respondents across different public sector fields, professions, and responsibilities. Second, all the decisions involved the use of tax money and had an impact on the public, whether large communities or specific categories of citizens (e.g., Dunleavy 1992; Niskanen 1971; Wilson 1989).

Our survey also measured subjects' managerial status (i.e., whether participants were in charge of managing subordinates), age, gender, educational background, and public sector field of employment.

Table 2: Survey Wave and Demographic Characteristics of Public Managers and Employees, by Experiment

Cognitive bias	Framing					Anchoring	Proportion dominance	Status quo		Asymmetric dominance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Experiment										
Wave 1: $N = 102$	X			X		X	X			X
Wave 2: $N = 102$		X	X	X		X	X	X		X
Wave 3: $N = 100$		X	X		X	X	X	X		X
Wave 4: $N = 196$		X	X		X	X	X	X	X	X
Wave 5: $N = 100$		X	X		X	X	X	X	X	X
Average age (years)	46	44	44	46	43	44	44	44	43	44
% Managers	41	67	70	43	75	66	64	69	66	64
% Females	75	53	51	69	49	54	56	52	46	56
% Scientific degree	39	33	40	38	37	38	38	37	36	38
% Humanistic degree	40	41	34	41	37	37	38	38	35	38
% Health care	15	17	16	16	16	17	16	17	17	16
% Education	73	45	43	67	39	46	49	44	37	49
% Administration	7	24	23	7	28	22	21	24	29	21

Note: Experiments 2 and 3 were included in waves 2 through 3 and were alternative to each other. In other words, participants randomly assigned to respond to experiment 2 did not see experiment 3 and vice versa. We dropped 55 responses in experiment 6 because of a typo in the text of the low anchoring condition in survey wave 1 that we remedied for in the subsequent launches.

Following recent practices in experimental work in behavioral public administration (e.g., Baekgaard 2017; Jilke, Van Ryzin, and Van de Valle 2016; Marvel 2016), our data come from a five-wave experimental survey of Italian public employees and managers. The five waves took place in June and July 2016 and were administered through the Qualtrics Company; the number of subjects who participated in each launch ranged from 100 to 196. To qualify for our study, subjects had to meet two criteria: currently working in Italy and working for the public sector. During each of the five waves of data collection, we ran a subset of the 10 experiments, as shown in table 2. In each of the launches, experiments were presented to subjects in random order.

1.3 Experimental Evidence

Table 2 shows the demographics of the pooled sample for each of the 10 randomized controlled trials. For each of the 10 experiments, we ran a series of two-sample comparison of means tests and chi-squared tests to investigate whether experimental groups differed across demographic characteristics. Within each randomized controlled trial, the experimental groups were not statistically different in terms of

managerial status, average age, gender, education, or field of employment, except for the following differences. In experiment 1, participants in the positive framing group and the negative framing group differed in terms of field of employment ($p = .029$). In experiment 3, subjects in the two experimental arms differed marginally with regard to academic background ($p = .069$) and differed significantly in terms of field of employment ($p = .012$). The percentage of managers was marginally higher in the positive framing group than in the negative framing group ($p = .059$) in experiment 4, and it was marginally lower in the high-anchor group compared with the low-anchor group ($p = .090$) in experiment 6. In experiment 8, public workers in the two experimental arms were marginally different in terms of educational background ($p = .066$). We did not detect any other differences.

Framing

Figure 1 shows the proportion of participants in experiment 1 who preferred the policy with the certain outcome over the policy with the probabilistic outcome, by experimental treatment. The proportion of risk-averse civil servants was higher under a positive frame (.82, $N = 44$) than under a negative frame (.33, $N = 58$), $p < .001$. Results of a logistic regression indicated that the odds of choosing the sure thing over the gamble were 9.24 times greater under a positive frame compared with a negative frame ($p < .001$). Figure 1 further displays the proportion of respondents in experiment 2 who selected the certain option instead of the probabilistic one in the two experimental groups. As expected, we observed risk-averse behaviors more frequently when civil servants were exposed to positively framed (.83, $N = 128$) rather than negatively framed (.21, $N = 118$) policy options, $p < .001$. Results of a logistic regression indicated that public employees presented with positively framed options were 17.92 times more likely to choose the sure thing over the gamble, relative to their counterparts in the negative frame condition ($p < .001$). Lastly,

Figure 1: Proportion of Subjects Preferring the Certain Option over the Probabilistic Option, by Framing (Experiments 1, 2, and 3)

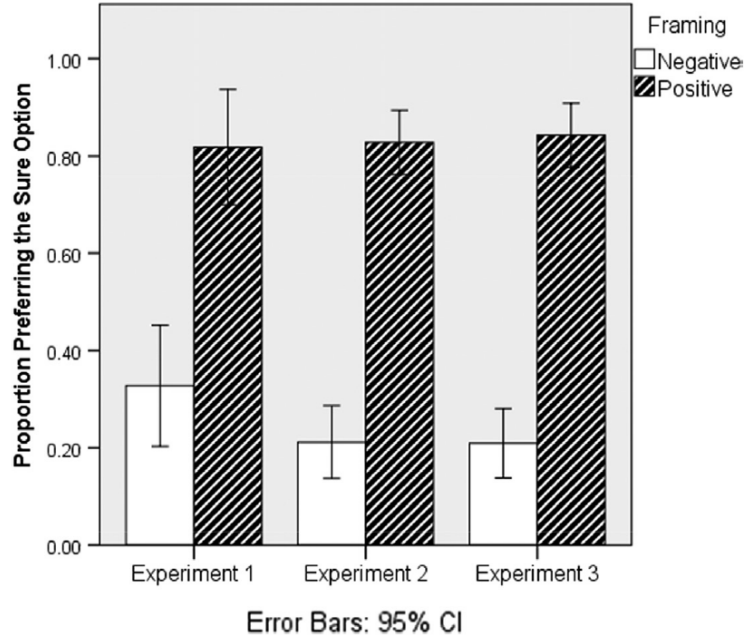
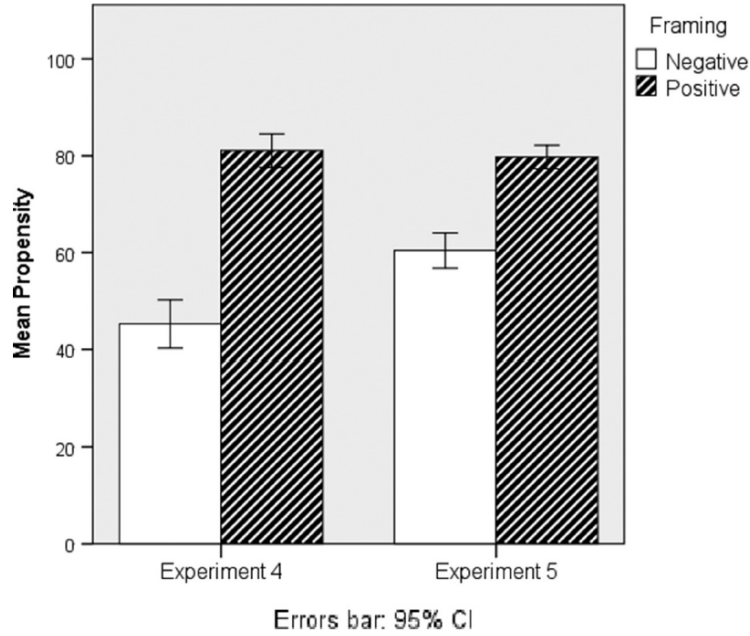


figure 1 shows the proportion of risk-averse participants in the two experimental conditions of experiment 3. As predicted, more public employees preferred the program with the sure outcome over the program with the probabilistic outcome in the positive framing group (.84, $N = 121$) compared with the negative framing group (.21, $N = 129$), $p < .001$. A logistic regression revealed that the odds of choosing the sure thing over the gamble were 20.21 times greater under positive framing relative to negative framing, $p < .001$.

Figure 2 shows the average propensity, on a 0–100 scale, to purchase the Attendance software in experiment 4 and 5, by experimental intervention. In experiment 4, civil servants who read that 80 percent of families and students were satisfied with the software (i.e., positive framing, $N = 105$) tended to report a higher propensity to purchase (81.06) compared with their peers who read that 20 percent of families and students were dissatisfied (i.e., negative framing, $N = 99$, 45.24), $p < .001$.

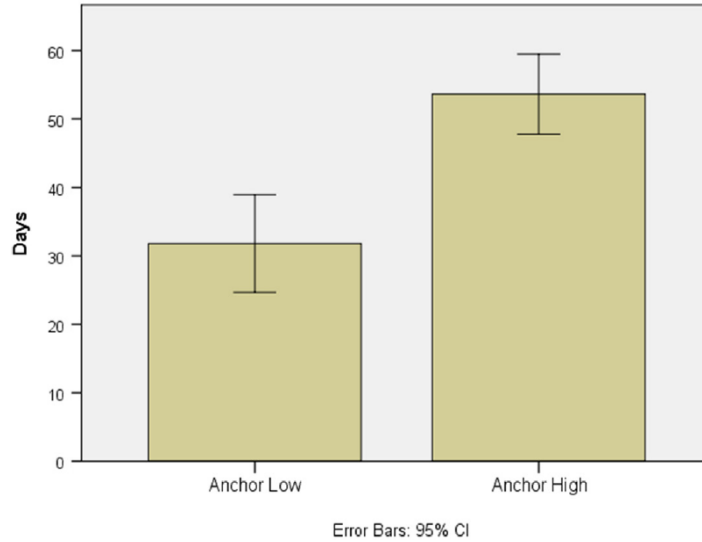
Figure 2: Propensity to Purchase the Software Attendance, by Framing (Experiments 4 and 5)



In experiment 5, we made it explicit that customers had to choose between two options: satisfied or dissatisfied with the Attendance software. Respondents were more willing to buy the software when they were exposed to the satisfaction rate (i.e., positive framing, $N = 189$, 79.72) than when they were exposed to the same information expressed in terms of dissatisfaction rate (i.e., negative framing, $N = 207$, 60.43), $p < .001$.

An analysis of variance (ANOVA) conducted on the pooled sample that included all participants in experiments 4 and 5 revealed that the framing effect was weaker in the latter case, that is, when we clarified that customers were given only two response options—satisfied and dissatisfied—compared with when we did not, $p < .001$. In other words, part of the framing effect disappeared after removing ambiguity about the number of satisfaction categories, thus making the positively and negatively framed pieces of information unambiguously equivalent.

Figure 3: Average Value of the Maximum Number of Days to Reply to Citizen Inquiries, by Anchor (Experiment 6)



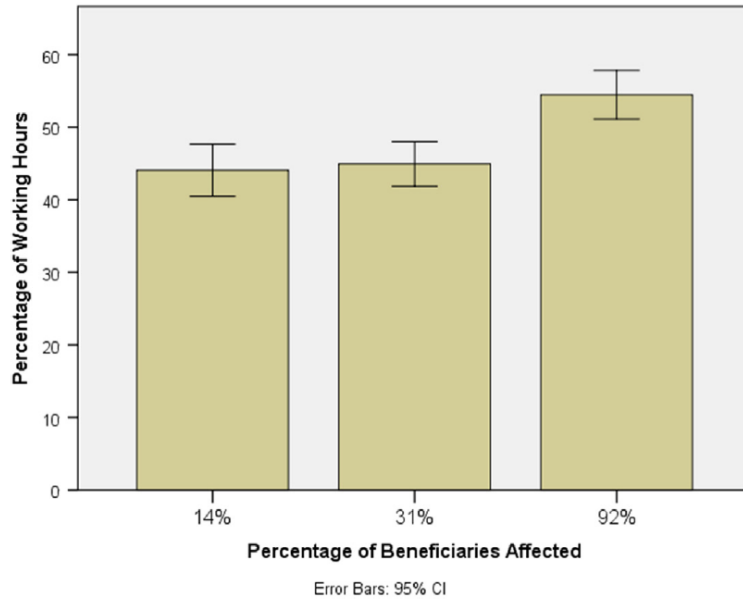
Anchoring

Figure 3 shows the sample average of the maximum number of days within which respondents in experiment 6 thought a public employee should reply to inquiries from citizens, separately for the two groups. On average, the maximum response time indicated by subjects in the low-anchor condition ($N = 228$; $M = 31.82$ days; $SD = 54.81$) was lower than the maximum response time indicated by respondents in the high-anchor condition ($N = 317$; $M = 53.62$ days; $SD = 53.07$), $p < .001$.

Proportion Dominance

A one-way ANOVA showed that the time participants were willing to spend for the selection of a new supplier significantly varied across the three levels of our experimental manipulation ($p < .001$). More specifically, figure 4 shows that the percentage of working hours that respondents in experiment 7 would dedicate to the selection of a new supplier was higher among subjects in the 92 percent level

Figure 4: Percentage of Working Hours That Participants Would Dedicate to Selecting a Contractor, by the Percentage of Beneficiaries Affected by the Service (Experiment 7)

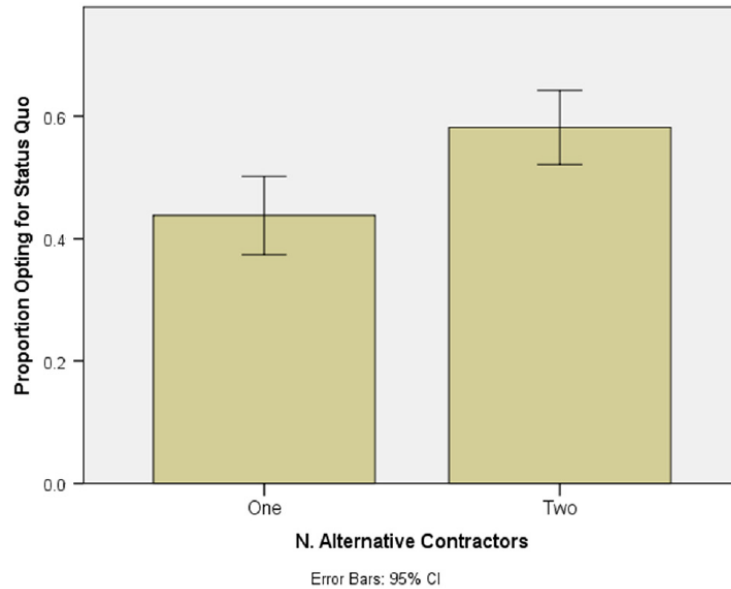


($N = 202$; $M = 54.48$; $SD = 24.26$) than for participants in both the 31 percent level condition ($N = 215$; $M = 44.94$; $SD = 22.88$; $p < .001$) and the 14 percent level condition ($N = 183$; $M = 44.09$; $SD = 24.64$; $p < .001$). The outcome was not significantly different between subjects in the 31 percent level group and subjects in the 14 percent level group, $p = .722$.

Status Quo

Figure 5 shows the proportion of respondents opting for the status quo in experiment 8 by experimental interventions. The results of a logistic regression revealed that the odds of sticking to the status quo were 1.78 times higher for participants presented with two alternative contractors ($N = 263$) rather than one ($N = 235$). Figure 6 displays the proportion of subjects reporting that they would publish the invitation to bid without trying any software fixes by experimental intervention

Figure 5: Proportion of Participants Sticking with Current Contractor (Status Quo), by Number of Alternative Contractors Available (Experiment 8)



(experiment 9). As predicted by the multiple alternatives effect, findings from a logistic regression revealed that the odds of sticking to the status quo were 5.74 times greater for respondents in the two-alternatives condition ($N = 141$) compared with respondents in the one-alternative condition ($N = 155$).

Asymmetric Dominance

Figure 7 reports the proportion of public workers opting for the print and web subscription in the treated and control group in experiment 10. A logistic regression showed that the odds of choosing the combo option were 1.59 times greater among participants presented with the decoy option (i.e., print subscription, $N = 302$) compared with participants who were not presented with the inferior option ($N = 298$).

Figure 6: Proportion of Participants Who Would Publish the Invitation to Bid without Trying Software Fixes, by Number of Alternative Software Fixes Available (Experiment 9)

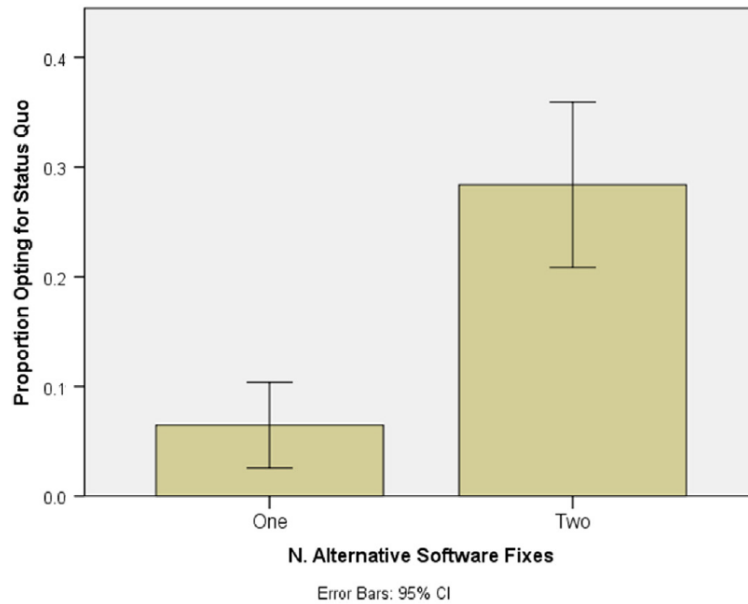
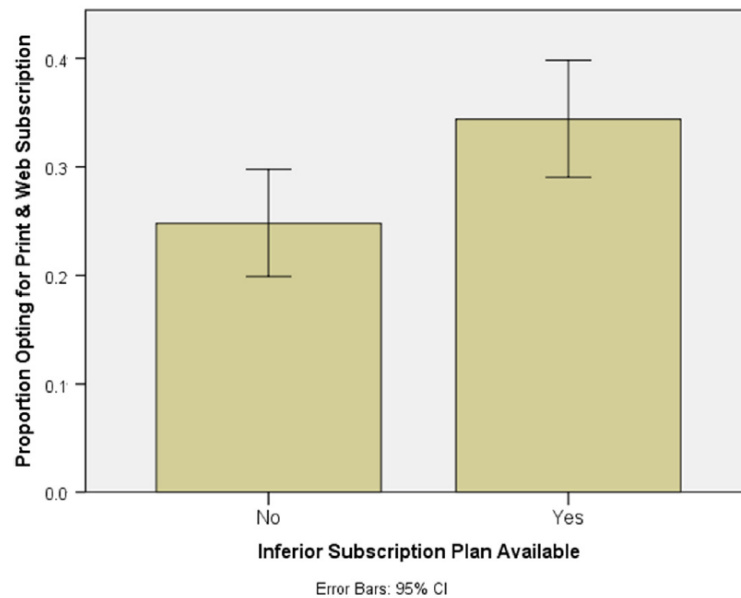


Figure 7: Proportion of Participants Opting for the Print and Web Subscription, with and without an Inferior Subscription Plan Made Available (Experiment 10)



1.4 General Discussion and Conclusion

To the best of our knowledge, this study represents one of the first attempts to investigate the effects of a broad range of cognitive biases on public workers' decisions across several policy areas and managerial tasks. We did so by conducting 10 randomized controlled trials on 600 workers from different public sector fields and with different managerial responsibilities. Subjects responded to an experimental survey that posed hypothetical questions about realistic work-related situations.

Results from experiments 1 through 3, which replicated Tversky and Kahneman's (1981) Asian disease experiment across several policy areas, revealed that public employees and managers were more risk averse when policy outcomes were formulated in terms of prospective gains and more risk seeking when equivalent outcomes were presented in terms of prospective losses.

Experiments 4 and 5 demonstrated that public sector workers may be more willing to purchase a good when equivalent information about previous customers' satisfaction is positively framed (i.e., percentage of clients satisfied with that good) rather than negatively framed (i.e., percentage of clients dissatisfied).

In experiment 6, subjects' judgments regarding the maximum time for response to citizens was heavily affected by the first piece of information presented to public employees in our sample, that is, the "anchor." In experiment 7, the proportion dominance effect increased participants' intentions to spend effort in an activity when the percentage of clients affected by that activity was higher, notwithstanding the fact that the absolute number of affected clients was identical across the percentage-level groups. In interpreting these results, which are consistent with previous scholarship about the proportion dominance effect (Erlandsson, Björklund, and Bäckström 2015), we cannot rule out the alternative hypothesis that subjects' behavior was not irrational but instead responded to a rationality that is different from maximizing the prosocial impact of a public manager's work. For instance,

drawing on the arguments of Stanovich and West (2000), if public sector workers in our sample were primarily preoccupied with gaining greater consensus and support from citizens, then it might have been rational for them to dedicate more time to activities affecting a larger share of clients—which we manipulated experimentally, unbeknownst to respondents.

The presence of multiple alternatives made respondents in experiments 8 and 9 more willing to stick to the current public administration contractor (i.e., status quo) when the number of alternative options increased. In deciding a journal subscription option to purchase on behalf of their public organization, subjects preferred the combo option when an asymmetrically dominated (or decoy) option was added to their choice set (experiment 10).

We fully acknowledge that our work is not without limitations and that our results should be interpreted accordingly. As far as the theoretical framework is concerned, we adopted a descriptive approach. Descriptive models are well equipped to observe whether there is an effect and its magnitude (i.e., whether the outcome variables systematically differ across experimental conditions and by how much), but they are poorly equipped to uncover the mechanisms underlying the effect. As far as the research design is concerned, the same general limitations that affect most survey experiments (e.g., Bouwman and Grimmelikhuijsen 2016; Harrison and List 2004) also influence our work. For instance, although our experimental research design scored high on external validity relative to laboratory experiments, there may be legitimate concerns about whether and to what extent our findings generalize to naturally occurring settings. Natural field experiments, which we strongly encourage scholars in our field to undertake, would allow better generalizability of results.

On a related note, the pattern of results that we observed in our sample, which consisted of only Italian public sector workers, may vary across different types of

units, treatments, operations, and settings (Shadish, Cook, and Campbell 2002). Moreover, although we tried to be as comprehensive as possible with regard to the typology of cognitive biases, our selection might have left out other systematic errors that have the potential to impinge on rational public policy and management.

Following the steps of scholars in other disciplines, we aim at making a twofold contribution to the advancement of research into public policy and management decision making. First, to the best of our knowledge, this is one of the first studies to investigate how cognitive biases impinge on the decisions made by public sector workers. Previous research has primarily targeted the general population of citizens without a specific focus on public policy and management decisions (e.g., Andersen and Hjortskov 2015; Baekgaard and Serritzlew 2016; Barrows et al. 2016; Marvel 2015, 2016; Moynihan and Lavertu 2012; Olsen 2015, 2017). Second, we studied a broad range of systematic decision-making errors. Overall, we echo work in other disciplines that comprehensively explores multiple cognitive biases affecting specific professions, such as judges (e.g., Guthrie, Rachlinski, and Wistrich 2001) and health care professionals (e.g., Blumenthal-Barby and Krieger 2015), thus providing a test of the external validity of their findings.

Our results have relevant implications for public policy making and the management of public organizations. First, the framing of policy options may significantly affect policy making. Policy makers should be aware of the human tendency to prefer a sure thing over a gamble when policy options are framed as prospective gains and to prefer a gamble over a sure thing when policy options are framed as prospective losses. This framing effect can occur in all situations in which policy makers have to choose between alternative courses of action based on their estimated outcomes. Having those pieces of information framed in a positive rather than a negative manner may lead to significantly different choices, which, in turn, have consequences for the society at large. Possible examples may span from

the choice between alternative vaccination programs, to environmental standards, transportation infrastructures, and fiscal interventions.

As for public organizations and their managers, our results demonstrate that managerial decisions may be highly dependent on systematic patterns of deviation from rationality. In particular, the propensity to purchase a management software among public managers and employees in our sample changed when subjects were presented with the percentage of clients satisfied rather than the percentage of clients dissatisfied with the software. Similar results can be expected anytime managers have to base their decisions on information that can be framed equivalently in positive or negative terms. Also, decisions on standards for responsiveness to citizens' inquiries turned out to be highly sensitive to the numeric anchors presented to respondents. This should raise awareness among public managers and employees that the initial piece of information offered to them when making decisions can significantly bias their subsequent judgments.

Furthermore, the intention to expend effort on an activity was conditional on the percentage rather than the absolute number, which was actually held constant, of clients affected by that activity. This finding should serve as a cautionary tale for public managers and employees, who routinely make decisions about the time to allocate to different tasks within their job responsibilities in the context of public organizations that benefit others. Then, preferences of the sample of participants in our study for a current public administration contractor and a managerial software in use in a hypothetical public organization varied, depending on the number of viable alternatives that we presented to them. This evidence speaks to any situations in which public managers have to choose between the status quo and new courses of action; in these situations, managers should ask themselves whether and how the presence of multiple alternatives is affecting their decisions.

Lastly, the selection of a journal subscription plan made by public sector workers

in our sample was consistently biased toward one of two plans when an inferior option was added to the choice set. Evidence of the asymmetric dominance effect should encourage public managers to carefully analyze whether a decoy is present among the options at their disposal and consider how this dominated alternative can systematically influence their decisions.

Our work highlights sizeable deviations from rational decision-making among public managers and employees. This speaks to the importance of taking these deviations into account while designing policy interventions, management systems and procedures. Additional experimental work might replicate these findings in different contexts and test similar treatments on more elaborate public decision scenarios. Future research is needed in related areas as well.

First, scholars in our field should engage in experimental work on the impact of de-biasing strategies aimed at mitigating the consequences of cognitive biases in the context of public policy and management decisions (e.g., Larrick 2004). This seems imperative in light of Herbert Simon's early argument that one of the main functions of government organizations is to cope with the limits of their members' "abilities to comprehend and compute in the face of complexity and uncertainty" (1978, 345). For instance, scholars in our discipline might run randomized controlled trials to test whether providing subjects with both positively and negatively worded information may mitigate the framing effect. Similarly, future research is needed that tests whether providing training and information about specific cognitive biases may reduce the risk of falling into those predictable traps.

Second, behavioral public administration researchers should further investigate the effect of re-biasing interventions that entail the use of one cognitive bias to offset another (e.g., Thaler and Benartzi 2001). More broadly, additional research is needed on how to leverage the architecture of choices without limiting the options available and without altering economic incentives to encourage desired

behaviors (e.g., Thaler 2015; Thaler and Sunstein 2008). In this respect, nudge theory particularly seems to suit the public sector, where high-powered incentives are rarely available. Indeed, Thaler noted that the U.S. federal government, for example, has “adopted the idea of automatic enrollment in many domains including its health care program,” but also “the concepts of behavioral economics have received bipartisan support in Congress” (Katz 2017). Public organizations and their managers should also follow the steps of the United Kingdom’s Behavioural Insights Team: more governments around the world seem interested in creating units dedicated to improving policies and services through the use of behavioral sciences evidence (OECD 2017).

2 Framing Effects under Different Uses of Performance Information: An Experimental Study on Public Managers²

Performance measurement has been heralded by New Public Management reformers as a core element in the modernization of public sector organizations, translating into practice principles of managerialism, economic rationality, and results orientation (Hood 1991; Osborne and Gaebler 1992; Pollitt and Bouckaert 2011). As these principles became globally diffused and recognized as the new mantra for the public sector, both performance measurement practices and related empirical research gained momentum (Kroll 2014; Moynihan et al. 2011). Research looking at the implementation and effects of performance measurement systems increasingly pointed to the limited use of performance information and started to explore the conditions under which it is used (Kroll 2015). If a rich literature now exists on the factors that enable the use of such information (for reviews, see Kroll 2014, 2015) - that is, on *when* performance information is used and by *whom* - less consideration has been devoted to understanding *why* and *how* performance information is used. Therefore, performance information use still remains a “black box” that needs unpacking.

The use of performance information has often been seen as a monodimensional concept, with most research only investigating whether or not use occurs (see, e.g., Moynihan and Pandey 2010). This suggests that there is a need to further explore the implications of differences in the conditions under which information is used (e.g., Moynihan 2009; Speklé and Verbeeten 2014). Moreover, most authors so far have focused on understanding the drivers of use, rather than the *ways in*

²This paper was conducted in collaboration with Nicola Bellé, Mariafrancesca Sicilia, and Ileana Steccolini. It has been published in the *Public Administration Review*.

which performance information is processed by users. This indicates that more investigation may be needed to understand how users deal with performance information.

Interestingly, the few studies focusing on how information is processed in the public realm refer to either citizens or politicians and show that the use of performance information, far from being a purely rational and reflective process, is affected by, among other things, primes, frames, and prior beliefs (Andersen and Hjortskov 2015; Baekgaard and Serritzlew 2016; George et al. 2018; Nielsen and Baekgaard 2013; Nielsen and Moynihan 2017; Olsen 2015). Surprisingly, little is known about how *public managers* process performance information and engage with it when taking decisions related to their responsibilities and whether they process information differently in different situations of performance information use. Indeed, managerial responsibilities and tasks encompass many different uses of performance information, including decisions on the evaluation of employees; the distribution of incentives; promotions; the allocation of financial and nonfinancial resources to organizational units, services, and managers; efforts to be deployed in different tasks and goals; future goals and strategies; cutbacks; and the prioritization and reconfiguration of services. Each of these decisions requires that relevant information be selected, interpreted, and elaborated. How these microprocesses work under different tasks and situations, and their related accuracy, have wider relevance in that they affect public employees' behaviors and commitment, the resources available to organizational units and services, the functioning of and directions taken by public sector organizations, the provision of public services, and, ultimately, the responses to citizens' needs.

The aim of this article is to look at how the types of situations in which managers use performance information may bring about different cognitive processes, with different levels of accuracy. To do so, the article combines insights from the

public administration, accounting, and psychological literatures to offer a twofold contribution. First, it provides evidence that public managers process information differently *under different performance information use situations*, being more likely to be subject to framing effects under ex ante uses of performance information than ex post; second, it shows that asking managers to justify their decisions does not mitigate such effects.

The article is structured as follows: The following section provides theoretical discussion of micro-processes of performance information use, drawing on the public administration, accounting, and psychological literatures. The next sections present the experimental designs and report the results. The final section discusses the results and draws conclusions.

2.1 Performance Information Use in the Public Sector: Toward a Focus on Managerial Microprocesses

An increasing body of literature has explored performance measurement system adoption (Bouckaert 1993; Meekings 1995), implementation (Julnes and Holzer 2001), and managerial use (e.g., Van Dooren and Van De Walle 2008). To better understand this phenomenon, a number of studies have sought to identify the factors that explain performance information use (e.g., Ammons and Rivenbark 2008; Kroll 2014, 2015; Melkers and Willoughby 2005; Moynihan and Lavertu 2012; Moynihan and Pandey 2010; Moynihan, Pandey, and Wright, 2012; Taylor 2011).

The studies described here have helped clarify the distinction between adoption, implementation, and use of performance information and highlight the related drivers and, in some cases, their relationships with organizational performance. They have generally looked at performance information as either an input or an output of managerial processes. However, there is a relative paucity of research on the microprocesses by which managers engage with performance information,

elaborating, interpreting, and, in sum, using it. This is potentially relevant since the performance measurement movement has tended to suggest a view of performance information as “objective” and “neutral” support for decisions made by politicians, citizens, and public managers.

However, this idea of neutrality has been challenged on different grounds. Information can be used intentionally and even misrepresented for political or perverse reasons (Moynihan, Pandey, and Wright 2012). Moreover, the bounded-rationality literature has suggested that judgment is systematically affected by cognitive limitations and information availability and tractability and thus will depart from pure rationality (Rabin 1998; Simon 1955).

Interestingly, how these phenomena unfold in the public realm has been studied mainly with reference to politicians and citizens. Such studies have shown that politicians’ and citizens’ use of performance information is affected by the framing of information (Olsen 2015), priming and frames (Andersen and Hjortskov 2015), prior beliefs (Baekgaard and Serritzlew 2016), negativity bias (Nielsen and Baekgaard 2013; Nielsen and Moynihan 2017), and institutional isomorphism (George et al. 2018). For example, Olsen (2015) shows that presenting hospital satisfaction results to citizens as “dissatisfaction rates” leads to more negative assessments of services than relying on “satisfaction rates.” However, the negative response is reduced by professional experience and prior exposure to satisfaction rates. Andersen and Hjortskov (2015) show that the dual-process theories of reflective and intuitive thinking provide a more adequate interpretation of citizens’ use of information than the expectation-disconfirmation model. Baekgaard and Serritzlew (2016) show that even the interpretation of unambiguous performance information by citizens is influenced by prior beliefs. In their study on politicians, Nielsen and Moynihan (2016) show that negativity bias affects the use of performance data in judging leadership responsibility. George et al. (2018) highlight that coercive and

normative pressures influence politicians' performance information use, whereas mimetic pressures and negativity bias appear to have a more limited impact on it.

Less attention has been devoted to understanding the conditions and possible biases and frames that affect public managers' microprocesses of decision making. This study contributes novel experimental evidence to a nascent stream of research investigating managers' misrepresentation of performance information. For instance, Meier et al. (2015) show that school principals systematically overestimate the performance of their own schools, and perceptual performance is only weakly associated with official performance. Bellé, Cantarelli, and Belardinelli (2017) have shown that anchoring and halo effects systematically bias performance ratings of public sector managers and employees.

This is an interesting research area, as public managers have often been depicted as driven by economic rationality, abiding by professional norms, and being equipped with the expertise necessary to provide public services and run day-by-day administration (Schedler 2003). These managerial prerogatives are likely to influence how managers process performance information. Managers may be expected to process performance information in a rational way, relying on their position and expertise to make decisions, allocate resources, evaluate staff, and more generally address various micromanagement processes. However, similar to citizens and politicians, public managers' actual behaviors may be affected by biases and cognitive limitations in the elaboration of performance information. As past research has suggested that managerial use or nonuse of information depends on a number of contextual, organizational, task-related, individual factors, similarly, it may be expected that there are organizational or task-related conditions that also affect the microprocesses of use—that is, how information is interpreted and processed by managers. Public managers face different situations and tasks in which they are expected to use and process performance information. What is

not clear is whether such processes of elaboration are similar (e.g., in terms of thoughtfulness and accuracy) across different situations—that is, under different types of performance information use. Moreover, public managers are often asked to justify, explain, and account for their decisions, yet it is not clear whether these processes of being held accountable improve the accuracy of their decisions.

This article explores whether the type of performance information use and being asked to justify related decisions affect the microprocesses by which public managers interpret and use performance information (and, more specifically, their accuracy). To do so, it combines the public administration and accounting literature on performance information uses and the psychological theory of equivalence framing.

Microprocesses of Performance Information Use: Public Administration and Accounting Literature

Conceptual contributions in the public administration literature (Behn 2003; Moynihan 2009; Moynihan, Pandey, and Wright 2012; Van Dooren, Bouckaert, and Halligan 2010) generally point out that performance measurement systems may serve different purposes, including evaluating, controlling, budgeting, motivating, celebrating, learning, and improving (Behn 2003) or control, steering, learning, and accountability (Van Dooren, Bouckaert, and Halligan 2010). Interestingly, however, most empirical studies focus on a monodimensional view of performance information use. Among the few empirical studies considering a variety of performance information uses, Moynihan, Pandey, and Wright (2012) distinguish between purposeful and political uses; Speklé and Verbeeten (2014), drawing on the accounting literature, show how different types of performance information use have different relationships with organizational performance. More generally, the accounting literature (referring to private sector firms) has theoretically and empirically pointed to a diversity of uses of performance information (Henri 2006;

Simons 1990; Vandenbosch 1999).

It is outside the remit of this article to discuss them at length or to provide an exhaustive review of such literature. However, drawing on the most commonly adopted classifications, it is possible to identify two main roles for performance information in organizations. Ex ante, it can be used to support planning, explore new possibilities, foster learning, define goals to be pursued, and allocate resources and efforts. This use of performance information has been described as *decision facilitating* (Demski and Feltham 1976), aimed at focusing attention and energies, stimulating new ideas and initiatives (Henri 2006; Vandenbosch 1999), reducing ex ante uncertainty (Tiessen and Waterhouse 1983), revising decision-making beliefs (Baiman 1982), assisting in problem-solving (Simon 1954), and improving employees' knowledge, thereby enhancing their ability to make organizationally desirable judgments and better-informed action choices (Sprinkle 2003). Through the decision-facilitating use, managers decide how to allocate resources and efforts on the basis of past results and expectations about the future. Sprinkle (2003) argues that performance evaluation involved in the decision-facilitating use is intended to improve future performance. Generally, under this type of use, managers are expected to scan for relevant information, analyze it systematically, take an exploratory stance, and ponder different possibilities and opportunities. They may also feel responsible for the future effects of their choices and thus pay more attention during the decision-making process, as suggested by studies showing that personal involvement increases the degree of attention to information content and reduces the reliance on simple heuristics (e.g., Borgida and Howard-Pitney 1983; Chaiken 1980; Harkness, DeBono, and Borgida 1985; Petty, Cacioppo, and Goldman 1981; Showers and Cantor 1985).

Ex post, performance information is used for assessing past performance and appraising and rewarding employees and organizations (Baiman 1982; Demski

and Feltham 1976) through monitoring, measuring, and evaluating. This use of performance information has been described as *diagnostic* (Simons 1990) or *decision influencing* (Demski and Feltham 1976), whereby information is used to provide feedback and to monitor and reward the achievement of preestablished goals, with a focus on exceptions, mistakes, and negative variances (Sprinkle 2003).

With respect to ex ante and ex post uses of performance information by managers in the public realm, there is no cue or past research on how information is actually processed or whether there are differences in the ways in which managers approach information among uses. To explore this issue, it is useful to rely on the psychological theory of equivalence framing.

Framing Effects under Ex Ante and Ex Post Uses of Performance Information

The bounded-rationality literature suggests that judgment is systematically affected by cognitive limitations and by information availability and tractability and thus will depart from pure rationality (Rabin 1998; Simon 1955). Decision makers tend to rely on a limited number of heuristic principles, which reduces the complex tasks of assessing probabilities and predicting values to simpler judgmental operations (Tversky and Kahneman 1974). In particular, the framing literature suggests that in the act of interpreting and processing information, individuals may be influenced by how information is framed (Levin, Schneider, and Gaeth 1998; Tversky and Kahneman 1981, 1986). In general, equivalence framing effects occur when individuals respond in systematically different ways to different but objectively equivalent pieces of information that are framed differently (Levin, Schneider, and Gaeth 1998; Rabin 1998; Tversky and Kahneman 1981, 1986). Several studies in different fields show that describing situations in terms of success instead of failure rates affects evaluations and decisions, as positive framing leads to more favorable

evaluations than negative framing (Kühberger 1998; Levin, Schneider, and Gaeth 1998).

Framing effects may arise from a lack of attention, and they are expected to occur less frequently when people think more carefully about the choice they are making (LeBoeuf and Shafir 2003; Sieck and Yates 1997; Simon, Fagley, and Halleran 2004; Smith and Levin 1996). Looking at the presence of framing effects will thus signal the type of cognitive processes that underlie the adoption of managerial decisions. In this article, framing effects are used to explore the characteristics of such processes, in terms of accuracy and thoughtfulness, under ex ante and ex post uses of performance information.

As pointed out earlier, when performance information is used ex ante to decide on future actions to be taken or resources to be allocated, managers tend to explore possibilities and alternatives, probably including more elements and paying more attention in their decisions. This may reduce their reliance on simple heuristics, translating into a more thoughtful and accurate processing of information. This suggests that equivalence-framing effects may be less likely as managers are able to evaluate all the information available, consider the consequences of their choices, and be more careful about the underlying phenomena described by the measures they are considering. When performance information is used ex post—that is, to evaluate and give feedback—there may be an expectation that managers will be more likely to narrowly focus on the specific task to assess performance and on the specific data provided to them, feel less personally involved in the outcome of their decision, and be less likely to include more elements in their decisions, thus relying more strongly on simple heuristics. They may thus be expected to be less accurate and thoughtful in their decision and more prone to be influenced by a framing effect. From these considerations, the following hypothesis arises:

Hypothesis 1: *Framing effects will be stronger under ex post uses of performance*

information than under ex ante uses.

Framing Effects, Performance Information Use, and the Justification of Decisions

The framing literature also provides useful lenses through which to explore whether asking managers to justify decisions helps improve decision making. Public managers are often asked to supply accompanying reports or qualitative explanations, which are aimed at contextualizing and justifying the decisions made in situations such as the allocation of resources and efforts or performance evaluations. Such reports are intended, on the one hand, to provide explanations and, on the other, to hold decision-makers accountable. It is thus interesting to explore whether these accountability requirements contribute to improve the accuracy and thoughtfulness of the way in which performance information is processed.

The literature shows that encouraging people to justify their choices—for example, explaining them in writing—can improve decisions by strengthening thoughtfulness (Davis and Bobko 1986; Fagley and Miller 1991; Koriat, Lichtenstein, and Fischhoff 1980; Rabin 1998; Sieck and Yates 1997; Stinessen 1985). Justifying one's reasoning is expected to “lead to greater thought about the choice, and hence less contamination by biasing factors such as framing” (Smith and Levin 1996, 284), as framing effects are less likely when the decision is processed to a greater degree (LeBoeuf and Shafir 2003; Simon, Fagley, and Halleran 2004). Studies looking at the moderating role of justifications provide mixed results. A number of studies, including Fagley and Miller (1991), Sieck and Yates (1997), and Takemura (1994), provide evidence of reduced framing effects when justification for decisions is required, suggesting that being asked to justify a decision encourages more thoughtful deliberation.

However, the persistence of framing effects in spite of a request to justify one's choices was found in other studies (LeBoeuf and Shafir 2003; Levin and Chapman

1990; Schooler and Melcher 1995; Schooler, Ohlsson, and Brooks 1993; Takemura 1993; Wilson and Schooler 1991). These last results may be explained by drawing on the cognitive dissonance literature (Festinger 1957), which suggests that individuals who adopt decisions that are apparently in conflict with their beliefs may experience mental discomfort and seek to reduce such conflict. The need to reduce cognitive dissonance may cause confirmatory bias (Rabin 1998, 31), whereby prior beliefs and opinions affect decisions. Along these lines, once people have taken a decision, they may want to provide support for it instead of changing their choice, and thus requesting justification of actions may not produce more accurate decisions. Moreover, cognitive dissonance may focus people's attention on nonoptimal criteria when they are asked to make explicit the reasons underlying a certain choice (Wilson and Schooler 1991). As people are often unaware of the processes and criteria underlying their decisions, asking them to justify their decisions may not necessarily increase decision-making thoughtfulness but rather result in people trying to confirm their choice by giving plausible responses (Wilson and Schooler 1991, 182). From this follows:

Hypothesis 2: *Asking for justification for decisions will not necessarily increase public managers' accuracy in processing performance information.*

2.2 Methods

Empirical Setting

To explore whether and to what extent framing effects influence public managers' decisions, five artifactual survey experiments were conducted using an online survey. The participants were public managers working in Italian municipalities with more than 5,000 inhabitants.³ In Italy, municipalities have jurisdiction over a large and

³In contrast to larger municipalities, in municipalities with fewer than 5,000 inhabitants, managerial responsibilities are taken on by politicians sitting on the municipal executive board.

heterogeneous number of services, including social care and assistance, education, local transportation, urban planning and security, waste disposal, and commercial activities. As such, whereas Italian municipal public managers share the same hierarchical level, their responsibilities vary from human resource management to financial planning, requiring them to take decisions under different situations. For these reasons, they are well suited to participate in this study.

Data Collection

All available email addresses of Italian municipal public managers were retrieved from the relevant official websites (for a total of 9,437 addresses). The online surveys were administered between February and July 2016, and a further version of the survey experiment was administered between November and December 2017.

A total of 1,207 public managers working in 630 municipalities participated in the five artificial survey experiments: 306 in survey Alpha, 68 in survey Beta, 83 in survey Gamma, 95 in survey Delta, and 655 in survey Epsilon. We sent 4,882 invitations to participate in survey Alpha, 1,430 in survey Beta, 1,210 in survey Gamma, and 1,915 in survey Delta, with response rates of 6.3 percent, 4.8 percent, 6.9 percent, and 5.0 percent, respectively. Invitations to participate in survey Epsilon were sent to all 7,999 contacts who had not opened the survey in the previous administrations, and the response rate was 8.1 percent. These relatively low response rates reflect the difficulties generally encountered when surveying Italian public managers (Ditillo et al. 2015; Liguori, Sicilia, and Steccolini 2012). Although the low response rates may detract from external validity, they do not impinge on the internal validity of the findings (Shadish, Cook, and Campbell 2002).

Out of 7,982 Italian municipalities (as of April 2017), 2,435 have more than 5,000 inhabitants.

Randomization Procedure

In experiment Alpha, subjects were asked to imagine themselves as the director of the Sports and Culture Department of a fictitious Italian municipality, and then they were presented with information about the customer satisfaction ratings of the municipality's sports facilities. A random subgroup of participants received this piece of information in a negatively framed fashion—that is, it was framed in terms of the percentage of customers who were dissatisfied with the sports facilities. The same information was positively framed for the remaining respondents—that is, it was framed in terms of the percentage of customers who were satisfied.

Experiment Beta was the same as experiment Alpha, but in experiment Beta it was made explicit that customers had been asked to choose between two options only: satisfied or dissatisfied with the sports facilities' services. The rationale for this clarification was to make the positively framed information unambiguously equivalent to the negatively framed information.

Experiment Gamma was the same as experiment Alpha, but subjects in Experiment Gamma were asked to make the same four decisions as in experiment Alpha and then type a justification in a space provided below each of the four sliders.

Experiment Delta departed from experiment Gamma the same way as experiment Beta departed from experiment Alpha. In other words, in experiment Delta it was made explicit that customers had been asked to choose between two options only: satisfied or dissatisfied with the sports facilities' services.

In order to remove any doubt that these results were driven by the particular order of the presentation of decisions, experiment Epsilon was designed to replicate experiment Beta but with a random order of decisions. The English translation of the scenarios proposed to the subjects is included in Appendix A.

Independent Variable

The framing of information provided to the participants is the main independent variable, which was operationalized following Olsen (2015). Framing of information can be positive or negative. Like in Olsen, for the positive framing, the percentage of satisfied customers was randomly drawn from the following set: 75 percent, 80 percent, 85 percent, 90 percent, and 95 percent. For the negative framing, the percentage of dissatisfied customers was randomly drawn from the following set: 25 percent, 20 percent, 15 percent, 10 percent, and 5 percent.

Dependent Variables

After being presented with the customer (dis)satisfaction rating, subjects had to make four decisions. First, participants were asked to rate the performance of the sports facilities' director on a 0–100 continuous scale. Second, they rated the performance of the sports facilities on a 0–100 continuous scale. Third, respondents indicated on a 0–100 continuous scale the amount of effort they would expend to improve the sports facilities. Lastly, subjects had to set the sports facilities' budget for the next year—relative to the current year—on an 11-point discrete scale ranging from –50 percent to +50 percent. Subjects expressed their preferences by moving a slider on their computer screen. The decisions about rating the performance of the sports facilities and their director are examples of *ex post* uses of performance information. The decisions about effort intentions and budget allocations represent *ex ante* uses of performance information. In experiment Epsilon decisions were displayed in a random order.

Table 3: Means and Standard Deviations of the Outcome Variables by Framing, Experiment Alpha

Experiment Alpha	Negative Framing			Positive Framing			Δ
	<i>N</i>	Mean	SD	<i>N</i>	Mean	SD	
Director's rating	144	64.56	26.11	161	77.19	17.27	12.63***
Facilities' rating	144	65.15	25.12	161	76.68	18.07	11.54***
Job effort	145	62.48	29.32	161	58.44	31.93	-4.04
Budget change (%)	145	15.56	17.17	161	15.71	17.20	0.15

*Statistically significant at the 5% level; **at the 1% level; ***at the 0.1% level.

2.3 Results

Table 3 reports the tests for balance across subjects' demographic characteristics for the five experimental surveys, separately by framing condition. For each of the five experiments, groups with negatively framed and positively framed customer satisfaction ratings were balanced and did not differ at the .05 level with respect to age, gender, graduate degree, and number of people reporting to the respondent.

Table 3 shows the means and standard deviations of the four outcome variables in experiment Alpha, separately for participants presented with negatively framed and positively framed customer satisfaction ratings. Subjects in the negative framing condition gave a lower rating to the director of the sports facilities ($M = 64.56, SD = 26.11$) compared with respondents in the positive framing condition ($M = 77.19, SD = 12.63$), $p < .001$. Similarly, the performance appraisal of the sports facilities tended to be lower under negative framing ($M = 65.15, SD = 25.12$) than under positive framing ($M = 76.68, SD = 18.07$), $p < .001$. The framing of customer satisfaction ratings did not seem to affect either participants' job effort intentions ($p = .255$) or the budget they would allocate to the sports facilities next year ($p = .936$).

Table 4 shows the results of experiment Beta. The results for framing effect held when it was made explicit that customers had to choose between two satisfaction

Table 4: Means and Standard Deviations of the Outcome Variables by Framing, Experiment Beta

Experiment Beta	Negative Framing			Positive Framing			Δ
	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	
Director's rating	30	56.43	22.89	38	73.68	16.80	17.25***
Facilities' rating	30	57.47	18.59	38	72.61	21.00	15.14**
Job effort	30	64.4	23.99	38	63.89	30.82	-0.51
Budget change (%)	30	20.33	14.02	38	15.79	16.21	-4.54

*Statistically significant at the 5% level; **at the 1% level; ***at the 0.1% level.

levels (satisfied or dissatisfied). Subjects in the negative framing condition gave again a lower rating to the director of the sports facilities ($M = 56.43, SD = 22.89$) relative to respondents in the positive framing condition ($M = 73.68, SD = 16.80$), $p < .001$. Similarly, the performance appraisal of the sports facilities tended to be lower under negative framing ($M = 57.47, SD = 18.59$) than under positive framing ($M = 72.61, SD = 21.00$), $p < .01$. The framing of customer satisfaction ratings did not affect either participants' job effort intentions ($p = .941$) or the budget they would allocate to the sports facilities next year ($p = .228$).

Table 5 reports the main results from experiment Gamma, which were consistent overall with findings from experiment Alpha. In experiment Gamma, public managers in the negative framing condition gave a lower rating to the director of the sports facilities ($M = 64.00, SD = 25.43$) relative to subjects in the positive framing condition ($M = 75.70, SD = 11.70$), $p < .05$. Similarly, the performance appraisal of the sports facilities tended to be lower under negative framing ($M = 66.15, SD = 22.67$) than under positive framing ($M = 76.00, SD = 19.37$), $p < .05$. As in experiment Alpha, the framing of customer satisfaction ratings did not affect either participants' job effort intentions ($p = .218$) or the budget they would allocate to the sports facilities next year ($p = .789$).

In experiment Gamma, respondents were asked to provide a justification for each

Table 5: Means and Standard Deviations of the Outcome Variables by Framing, Experiment Gamma

Experiment Gamma	Negative Framing			Positive Framing			Δ
	<i>N</i>	Mean	SD	<i>N</i>	Mean	SD	
Director's rating	46	64.00	25.43	37	75.70	18.60	11.70*
Facilities' rating	46	66.15	22.67	37	76.00	19.37	9.85*
Job effort	46	61.70	29.22	37	69.95	31.20	8.25
Budget change (%)	46	16.30	20.91	37	15.14	18.20	-1.16

*Statistically significant at the 5% level; **at the 1% level; ***at the 0.1% level.

Table 6: Means and Standard Deviations of the Outcome Variables by Framing, for Public Managers Who Provided a Justification for Their Decisions in Experiment Gamma

Experiment Gamma	Negative Framing			Positive Framing			Δ
	<i>N</i>	Mean	SD	<i>N</i>	Mean	SD	
Director's rating	22	56.86	29.17	17	78.24	13.11	21.37**
Facilities' rating	21	62.24	28.24	16	76.13	18.30	13.89
Job effort	17	65.18	28.86	16	66.81	34.15	1.64
Budget change (%)	21	20.00	20.98	17	15.88	20.93	-4.12

*Statistically significant at the 5% level; **at the 1% level; ***at the 0.1% level.

of the four decisions they had to make. The pattern of results for respondents providing justifications for their choices (see table 6) did not significantly depart from results from the pooled sample (table 7). The size of the framing effect for the first two decisions was even larger among those providing a justification for their decisions ($M = 21.37, SD = 7.62$ and $M = 13.89, SD = 8.12$, respectively) than for those who did not ($M = 3.01, SD = 6.37$ and $M = 6.46, SD = 5.48$, respectively).

The results of experiment Gamma suggest that the framing effect was unaffected by the request to justify decisions. Contrary to expectations, asking public managers to justify their decisions did not protect them from framing bias. To the contrary, the data provide evidence that the size of framing bias occurring in the first decision was larger for public managers who justified it.

Table 7: Means and Standard Deviations of the Outcome Variables by Framing, for Public Managers Who Did Not Provide a Justification for Their Decisions in Experiment Gamma

Experiment Gamma	Negative Framing			Positive Framing			Δ
	<i>N</i>	Mean	SD	<i>N</i>	Mean	SD	
Director's rating	24	70.54	19.89	20	73.55	22.36	3.01
Facilities' rating	25	69.44	16.56	21	75.90	20.60	6.46
Job effort	29	59.66	29.74	21	72.33	29.39	12.68
Budget change (%)	25	13.20	20.76	20	14.50	16.05	1.30

*Statistically significant at the 5% level; **at the 1% level; ***at the 0.1% level.

Experiment Delta included both the explicit information about the fact that customers had to choose between two satisfaction levels (satisfied or dissatisfied) and the request to justify the decisions. Table 8 shows that the combination of these two variations did not affect the pattern of results observed in the other three experiments. The average rating of the sports facilities director was again lower in the negative framing condition ($M = 64.07, SD = 24.27$) relative to the average rating in the positive framing condition ($M = 75.83, SD = 11.76$), $p < .01$. Also, the performance appraisal of the sports facilities tended to be lower under negative framing ($M = 67.20, SD = 23.25$) than under positive framing ($M = 76.09, SD = 17.17$), $p < .05$. There was no detectable effect in either the participants' job effort intentions ($p = .885$) or the budget they would allocate to the sports facilities next year ($p = .616$).

The pattern of results for respondents who did provide justifications for their choices (see table 9) did not significantly depart from the pattern of results from the sample of public managers who did not provide justifications (see table 10).

Random order of decisions did not change the results. Table 11 reports the differences in responses by participants in experiment Epsilon. When the customer (dis)satisfaction rate was negatively framed, participants gave a lower rating to the

Table 8: Means and Standard Deviations of the Outcome Variables by Framing, Experiment Delta

Experiment Delta	Negative Framing			Positive Framing			Δ
	<i>N</i>	Mean	SD	<i>N</i>	Mean	SD	
Director's rating	41	64.07	24.27	54	75.83	15.67	11.76**
Facilities' rating	41	67.20	23.25	54	76.09	17.17	8.90*
Job effort	41	60.61	29.79	54	61.46	27.45	0.85
Budget change (%)	41	15.12	18.18	54	16.85	15.27	1.73

*Statistically significant at the 5% level; **at the 1% level; ***at the 0.1% level.

Table 9: Means and Standard Deviations of the Outcome Variables by Framing, for Public Managers Who Provided a Justification for Their Decisions in Experiment Delta

Experiment Delta	Negative Framing			Positive Framing			Δ
	<i>N</i>	Mean	SD	<i>N</i>	Mean	SD	
Director's rating	32	65.03	25.06	33	79.15	13.60	14.12**
Facilities' rating	30	71.97	21.44	33	81.45	12.22	9.49*
Job effort	28	61.86	31.59	30	59.17	29.42	-2.69
Budget change (%)	29	14.48	16.82	30	14.00	15.67	-0.48

*Statistically significant at the 5% level; **at the 1% level; ***at the 0.1% level.

Table 10: Means and Standard Deviations of the Outcome Variables by Framing, for Public Managers Who Did Not Provide a Justification for Their Decisions in Experiment Delta

Experiment Delta	Negative Framing			Positive Framing			Δ
	<i>N</i>	Mean	SD	<i>N</i>	Mean	SD	
Director's rating	9	60.67	22.26	21	70.62	17.55	9.95
Facilities' rating	11	54.18	23.97	21	67.67	20.50	13.48
Job effort	13	57.92	26.48	24	64.33	25.10	6.41
Budget change (%)	12	16.67	21.88	24	20.42	14.29	3.75

*Statistically significant at the 5% level; **at the 1% level; ***at the 0.1% level.

Table 11: Means and Standard Deviations of the Outcome Variables by Framing, Experiment Epsilon

Experiment Epsilon	Negative Framing			Positive Framing			Δ
	<i>N</i>	Mean	SD	<i>N</i>	Mean	SD	
Director's rating	318	60.11	25.06	301	68.84	22.83	8.73***
Facilities' rating	326	61.50	23.87	304	68.44	22.26	6.94***
Job effort	324	58.82	28.06	307	56.69	28.69	-2.13
Budget change (%)	330	17.53	2.86	308	18.70	3.12	-1.16

*Statistically significant at the 5% level; **at the 1% level; ***at the 0.1% level.

director of the sports facilities ($M = 60.11, SD = 25.06$) relative to their colleagues in the positive framing condition ($M = 68.84, SD = 22.83$), $p < .001$. The performance appraisal of the sports facilities was again lower under negative framing ($M = 61.50, SD = 23.87$) than under positive framing ($M = 68.44, SD = 22.26$), $p < .01$. The framing of customer satisfaction ratings did not affect participants' job effort intentions ($p = .347$) or the budget they would allocate to the sports facilities in the following year ($p = .380$).

In sum, subjects in the negative framing condition gave a lower rating to the director of the sports facilities and to the sport facilities themselves compared with respondents in the positive framing condition. By contrast, the framing of customer satisfaction ratings did not affect either participants' job effort intentions or the budget they would allocate to the sports facilities next year. The pattern of results was consistent across the five experiments. Because some of the subjects were located in the same municipality, the analysis was replicated using multilevel models to take into account the possible hierarchical structure of the data. Likelihood ratio tests indicated that the hierarchical structure of the data had no significant impact on estimates. Out of the 20 likelihood ratio tests (four per experiment), 15 indicated that a multilevel model was not a better fit to data compared with ordinary least squares. In the five cases to which likelihood ratio tests were significant, the results

did not change. In no case did a discrepancy emerge between the ordinary least squares estimates and the estimates from the multilevel model. Taken together, the findings suggest that the framing effects were contingent on the type of performance information use.

2.4 Discussion and Conclusions

New Public Management proposes performance measurement as a possible solution to improve decision making and performance in the public sector. However, while a number of studies have investigated when performance information is used and by whom, much less attention has been paid so far to investigate why and how performance information is used. This article helps unpack the performance measurement “black box” by exploring how public managers process information when faced with different tasks and related decisions. More specifically, integrating insights from public administration, accounting, and psychology, the article sheds new light on two related issues. First, it investigates whether public managers’ accuracy in processing information varies under different performance information use situations—that is, whether they are subject to stronger framing effects under ex post uses of performance information than under ex ante uses. Second, it explores whether the request for justification of decisions increases public managers’ accuracy in processing performance information.

The findings from five artifactual survey experiments appear to confirm that framing bias may be contingent on the *type of information use*, suggesting that public managers process information differently under different performance information use situations. In particular, when performance information is used ex post - that is, to evaluate organizations or individuals - the framing effect appears to hold. On the contrary, when information is used ex ante - that is, to decide on the allocation of financial resources or job efforts - the framing effect does not appear to

hold. This suggests that information is processed differently when it is used for ex post and ex ante purposes. When information is used to provide ex post feedback, public managers may be more likely to simply focus on their task as evaluators, taking into consideration fewer elements in their decisions and spending less effort and relying more strongly on simple heuristics. They may be expected to be less thoughtful and more sensitive to framing effect. On the contrary, when public managers' tasks refer to the allocation of resources and efforts, they will be more likely to scan the available information in an open-minded process, exploring and pondering possible alternatives and probably paying more attention to the available information while reducing reliance on simple heuristics. This will translate in their being less vulnerable to framing effects.

It must be recognized, however, that there may be other reasons behind the different impact of framing in the distinct decision scenarios. This possibility calls for further investigation, as more typologies of managerial decisions when ex ante and ex post uses of information are involved may be investigated.

The findings also show that, contrary to expectations, asking public managers to justify their decisions does not mitigate the framing effect, which seems to suggest that it does not increase the accuracy of the cognitive processes related to the use of performance information. If anything, in some cases the framing effect was higher among public managers who provided justifications. This suggests that the writing of reports and other forms of explanation and justification of decisions to keep public managers accountable may not be sufficient to encourage more in-depth analysis and exploration and thus reduce the risk of superficial decision making. This evidence is consistent with previous studies that found the persistence of framing effects even if when justification of decisions was elicited (Levin and Chapman 1990; LeBoeuf and Shafir 2003; Schooler and Melcher 1995; Schooler, Ohlsson, and Brooks 1993; Takemura 1994; Wilson and Schooler 1991).

These results may be attributable to cognitive mechanisms that encourage people to avoid questioning their previous decisions, such as confirmatory bias or other psychological mechanisms used to cope with cognitive dissonance (Festinger 1957), for example, attempts to justify choices by referring to criteria that are plausible but may not have necessarily been the bases for the actual decision (e.g., Wilson and Schooler 1991).

This article provides a twofold contribution. First, it is among the first studies to adopt an experimental approach to look at the processes by which performance information is used by public managers. Second, drawing on the insights of accounting literature, it sheds new light on the relevance of distinguishing among different types of performance information use, as different uses appear to bring about different cognitive processes and are subject to framing effects to different extents.

The results bear interesting implications for managers and reformers, suggesting that particular attention should be given to avoid framing effects especially in ex post uses of performance information, whereby they may affect the fairness of evaluation processes. This means that the way information is provided may be central in affecting the evaluation of public managers and even related promotion paths and incentives or the evaluation of organizational units' performance, including performance-related sanctions and rewards. On the other hand, they also suggest that asking managers to justify their decisions may not always be sufficient to ensure that they are accurate in their information processing. However, on a more positive note, the results also suggest that, at least when decisions about future courses of actions are to be taken and managers are asked to scan the environment to explore options and possibilities, they will be more open-minded and thus less prone to bias. Future research may explore further mechanisms that may mitigate the framing bias in public managers' decisions as well as investigate framing effects

under different decision situations.

As with any piece of research, this article has limitations. While this experimental design is better equipped to avoid internal validity concerns, some potential threats to external and ecological validity exist. First, using fictitious information may call into question the behavioral parallelism assumption, whereby there is no guarantee that in making decisions in a real context with real incentives, public managers would behave the same way that they did in these experiments. For example, as emphasized by one anonymous reviewer, *ex ante* decisions on resource and effort allocations involve several complex trade-offs that cannot be fully captured in this study. Second, the pattern of results observed with Italian local public managers may vary across different populations. On a related note, it cannot be excluded that municipal managers participating in this study might be affected by framing bias in a different manner compared with their counterparts who did not complete the survey. Finally, it may be possible that the findings were contingent on the specific decisions designed for these experiments. Replications of the study by means of survey, lab, field, and natural experiments adopting a between-subjects design to directly manipulate distinct *ex-post* and *ex-ante* decisions, across different contexts, units, and operations would allow for better generalizability of this piece of evidence (Bellé and Cantarelli 2017; Van Ryzin, Riccucci, and Li 2017; Walker, James, and Brewer 2017).

3 On iron cages and suboptimal choices: An experimental test of isomorphism microfoundations in the public sector.⁴

Understanding how and why organizational arrangements converge and practices spread has attracted sustained scholarly attention over the last decades. One stream of research, called neo-institutionalism, has focused on investigating the dynamics of such homogenization or isomorphism - from the Greek “isos” [identical] and “morphe” [form] - in a variety of empirical settings. In particular, public organizations have been investigated both as a source of and a context where institutional pressures unfold (Frumkin and Galaskiewicz 2004). On the one hand, scholarship on isomorphism has portrayed public organizations as the originators of formal and informal norms that would ultimately influence private organizations, resulting in a progressive standardization of the arrangements under analysis (Meyer and Rowan 1977; Fligstein 1990, 1991; Orru, Biggart, and Hamilton 1991; Vasudeva 2013). On the other hand, scholars have recognized that public organizations are extremely susceptible to institutional pressures (Frumkin and Galaskiewicz 2004).

The scope of the studies that adopt this perspective in the public administration literature is broad. It ranges, for example, from the spread of privatization in the telecommunication industry among the member countries of the Organization for Economic Cooperation and Development (Fink 2011) to policy transfer in the European Union (Radaelli 2000), and from branding initiatives among universities in the United States (Fay and Zavattaro 2016) to the wide acceptance of extravagant position-related consumption in local governments in China (Gong and Xiao 2017). These studies are not only disparate in terms of the geographical coverage,

⁴This paper was conducted in collaboration with Nicola Bellé, Paola Cantarelli, and Valentina Mele. It has been published in the *International Public Management Journal*.

jurisdiction, and policy domain, but more importantly, they exemplify a notion of homogenization that spans from institutional arrangements to policy issues, and from managerial practices to inappropriate behaviors of civil servants.

Notwithstanding differences in their unit and level of analysis, these studies have a common denominator: their analytical reliance on one or more of three archetypes of isomorphic pressures that lead to homogenization (Meyer and Rowan 1977; DiMaggio and Powell 1983); i.e., influence from higher-level agencies (i.e., *coercive* isomorphism), mimicry of successful peers as a strategy to cope with environmental uncertainty (i.e., *mimetic* isomorphism), and strong ties with professionalization (i.e., *normative* isomorphism).

Disentangling the three types of isomorphism is unanimously considered a difficult task, since they often overlap and co-occur (Lodge and Wegrich 2005; Villadsen 2011; Teodoro 2014). A review by Mizruchi and Fein (1999) found that out of 26 articles attempting to operationalize and empirically test the definitions of isomorphic pressures developed by DiMaggio and Powell (1983), only two succeeded in operationalizing all three.

Irrespective of whether one or more types of isomorphic pressures is at work, homogenization displays a mixed record in public administration literature when it comes to societal outcomes. Some studies suggest skepticism about the likelihood that isomorphic pressures—and the bandwagoning behaviors and ritual constraints associated with them—will systematically lead to positive solutions for society (Pollitt 2001; Frumkin and Galaskiewicz 2004; Ashworth, Boyne, and Delbridge 2007; Pina, Torres, and Royo 2010; Kallio and Kuoppakangas 2013; Ammons and Roenigk 2015; Gong and Xiao 2017). This seems to be in line with earlier literature suggesting that isomorphism may be harmful when presented as a perverse alternative to market pressures and maximizing behavior (DiMaggio and Powell 1983; Abrahamson 1991; Dacin 1997). Others seem to point to the spread of

arrangements or practices that are intrinsically positive, such as the fulfillment of public record requests (ben-Aaron et al. 2017), the deployment of diversity management procedures (Pitts et al. 2010), or the compliance with safety standards (Teodoro 2014). Even in these instances, however, scholars suggest that the positive outcome of isomorphism may result more from a search for legitimacy than from a purposeful intervention to improve public performance (Laegreid, Roness, and Rubecksen 2007).

In sum, isomorphism unleashes its explanatory potential, especially when we try to understand the spread of arrangements and practices that do not necessarily display technical superiority. However, few studies have done other than infer the micro level at which isomorphic forces kick in and influence individuals' decisions. This gap, identified and discussed by PA scholars (Grimmelikhuijsen et al. 2017:50), resonates with a broader debate in organization theory arguing that “despite early scholars' attention to micro level psychological and sociocognitive aspects of institutions [...], the organizational research of the last two decades has focused primarily on organizations and field-level units of analysis” (Bitektine and Haack 2015:49). In a similar vein, it has been contended that while “for almost two decades scholars have stressed the need to make the micro-foundations of institutional theory more explicit [...], curiously there has been limited progress in this effort” (Powell and Colyvas 2008:276). Therefore, adopting a micro-foundations perspective to investigate central constructs in neo-institutional theory like isomorphism remains a significant research opportunity (Felin, Foss, and Ployhart 2015).

The current study seeks to contribute to address this knowledge gap by locating isomorphic pressures at the individual level and by testing whether public sector workers make managerial decisions on the grounds of technical superiority or as the result of exposure to such pressures. In turn, isomorphic pressures are exerted by collective actors, such as governments or professional organizations, who act

upon some collective and socialized legitimacy judgments (Bitektine and Haack 2015). Therefore, we embrace a micro-foundations perspective that is focused on bringing individuals back in without attributing explanatory exclusivity to the micro level where their agency is enacted. In other words, we adopt an approach such that “micro-foundations are embedded in a larger conversation related to multilevel theorizing and empirics” (Felin et al. 2015:586).

Specifically, we conducted eight randomized controlled trials with 764 public employees to test whether and to what extent isomorphic pressures affect public workers’ decision making at the individual level. In addition, in order to gain a more refined understanding of the causal mechanisms driving our experimental results, we supplemented the analysis of quantitative data from our randomized controlled trials with the qualitative inquiry of interview data collected from a subsample of participants.

3.1 Theoretical Background and Hypotheses

Having noted that our study looks at the independent effects that coercive, mimetic, and normative isomorphic pressures (DiMaggio and Powell 1983, 1991) have on public employees’ managerial decisions, we now turn to the definition and discussion of those pressures offered by the literature, which has substantially maintained and extended, more than challenged, the taxonomy proposed in the work of DiMaggio and Powell (e.g., Deephouse 1996; Glynn and Abzug 2002; Williamson and Cable 2003; Frumkin and Galaskiewicz 2004; Lodge and Wegrich 2005; Bellé 2010; Fay and Zavattaro 2016; Gong and Xiao 2017).

Coercive isomorphism suggests that homogenization occurs through the formal and informal pressures that a superordinate organization exerts on a subordinate organization. The resource-based view articulates dependency as the material dependence of certain organizations that will conform to the expectations of others

in order to secure inputs (Bovaird and Downe 2006). Dependency may also be conceptualized as power, especially of a political nature, which determines the course of action of the more vulnerable organizations (Radaelli 2000; Fink 2011; Gong and Xiao 2017). The enactment of such pressure ranges from very formal, such as legal requirements or health and safety regulations (Dacin 1997), to contractual obligations with other actors (Ashworth et al. 2007), to more subtle forms of imposing an organizational model on a dependent organization, originating in the exchange relationship (Currie and Suhomlinova 2006). One of these forms is the symbolic effect of regulation (Deephouse 1996), so that the pressures at work may be moderately binding and actually signal to participants the legitimacy of the prescribed behavior, consistent with the notion that “many myths also have legitimacy based on legal mandates” (Meyer and Rowan 1977:148). As such, coercive isomorphism has been expected to play an important role in public organizations (Frumkin and Galaskiewicz 2004), even more so in centralized public sectors (Meyer, Scott, and Strang 1987); i.e., those in which all local organizations across a nation depend on a central authority, and regional autonomy is scarce or not allowed (Lodge and Wegrich 2005).

A promising theoretical development is to look at the interplay between bureaucratization and coercive pressures. Following bureaucratic theory, with its emphasis on technocratic rationality, we may expect public organizations to follow and therefore converge towards one technically best solution. And yet, “the role of coercive forces in institutional theory highlights the impact of political rather than technical influences on organizational change” (Ashworth et al. 2007:167). Moreover, coercive pressure towards standardization is not the only available option. The literature has shown that formalization and centralization, two key attributes of bureaucratization, result from isomorphic pressures; for example, when public organizations become subject to oversight by higher jurisdictions. Building on

these premises, we formulated and tested Hypothesis 1.

Hypothesis 1: *Given options A and B, a coercive isomorphic pressure towards A increases the probability that public employees will choose A, even if A is inferior to B.*

We next turn to mimetic isomorphism, which refers to the process through which organizations emulate the arrangements or practices of others (Tolbert and Zucker 1983; Haunschild and Miner 1997). They do so typically in contexts of uncertainty that include doubts about the environmental conditions, goal ambiguity, and poorly understood organizational technologies (DiMaggio and Powell 1991; Deephouse 1996; Bovaird and Downe 2006; Currie and Suhomlinova 2006). Mimetic learning may occur directly, when exposure to and contacts with organizations introduce new ideas, and indirectly, when personnel hired from those organizations bring in fresh ideas (Ammons and Roenigk 2015) and reproduce the procedures developed in the previous organizational setting (Compagni, Mele, and Ravasi 2015). It may also be the result of an intentional search for solutions by an organization that turns to peers, especially well-performing ones (Haveman 1993; Gimeno et al. 2005), for ideas or suggestions on how to handle a policy decision or other concern (Fink 2011; Fay and Zavattaro 2016; ben-Aaron et al. 2017). In this case, it encourages and motivates organizations to learn from each other (Gong and Xiao 2017). Mimetic isomorphism poses serious challenges to the conventional notion of public action as the pursuit of actual performance improvements.

In a nutshell, we would expect public organizations to emulate and apply the successful arrangements or practices of other organizations to their own context. For example, initiatives designed to facilitate the spread of best practices, such as benchmarking and awarding schemes (Borins 2000; Hartley and Downe 2007; Ammons and Roenigk 2015), are among the institutional devices that facilitate mimetic pressures. Trying to learn from peers may be a rule of thumb developed by

individuals to cope with uncertainty and imperfect information (Gigerenzer et al. 1999; Artinger et al. 2015). Like all of the heuristics, in some situations this may lead to good decisions, especially when the peers are well-performing. However, wide evidence exists that heuristics may lead to cognitive biases and systematic departures from rational decisions (Tversky and Kahneman 1974; Kahneman 2000). More precisely, at least three main problems arise in reality when organizations try to learn from others. One is that organizations often import best practices without exerting due diligence on the applicability of such solutions to their context. Without adjustment and recalibration, it might be that what has proven effective in one specific setting may not be as successful when exported. Second, organizations facing severe institutional stress may not even look for best practices, but simply rely on cognitive shortcuts (Fink 2011) by turning to those they perceive as the most successful organizations and embracing their practices. Third and more importantly, a successful practice often is imbued with perceived legitimacy or appropriateness without any full evaluation of all options and their potential impact (Lodge and Wegrich 2005). In this vein, a caveat has been offered by several scholars that mimetic forces are behind the widespread adoption of management practices, driven more by fads and fashions than by empirical evidence of performance benefits (Abrahamson 1991; Kieser 1997; Ashworth et al. 2007; Fay and Zavattaro 2016). Therefore, based on these previous works, we formulated and tested Hypothesis 2.

Hypothesis 2: *Given options A and B, a mimetic isomorphic pressure towards A increases the probability that public employees will choose A, even if A is inferior to B.*

In our opinion, however, the strong symbolic valence attached to the devices mentioned earlier, together with the active role of professional communities in processing information and in designating the best solutions based on codified standards, puts them at the junction between mimetic pressures and the last type

of isomorphism we will present next. Normative isomorphism relies heavily on the notion of professionalization (DiMaggio and Powell 1991). This can be conceived as a centripetal force, resulting from a shared curriculum (Palmer, Jennings, and Zhou 1993) and certification processes often required to access specific government positions (Lodge and Wegrich 2005). It has also been studied as the explicit outcome of the activities of professional associations that lend their legitimacy and cognitive support to maintain stability or trigger change of specific professional standards (Greenwood, Suddaby, and Hinings 2002) and communities. Once established, tied professional networks may both span and constrain jurisdictional boundaries, as often is the case in professional bureaucracies (Hood 2000) or epistemic communities (Haas 1992). Whether or not such professional rules are formalized, they are prescriptive in nature (Scott 1995).

Professional norms abound in contexts and processes dominated by expertise and technocratic reasoning (Fink 2011) where, if successful, they get a cognitive upgrade and assume the taken-for-granted character that secures them a smooth enforcement. In other words, education, socialization, and filtering bolster convergence in individual orientations, which in turn bolster behavioral convergence among members of the same profession and occupants of similar positions across public organizations (Bovaird and Downe 2006; Currie and Suhomlinova 2006; Pitts et al. 2010; Teodoro 2014). Professional socialization may affect public managers' perception of what is a good policy (Vasudeva 2013), drawing on the experience of exemplary organizations, distilling and codifying expert knowledge, and channeling it through conferences and publications (Ammons and Roenigk 2015). A more ambiguous normative pressure exerted by professions is linked with career opportunities. In order to pursue career advancement, public managers may end up making decisions consistent with professional norms but not based on their technical superiority, even when they conflict with the directives attached to their

organizational roles (Teodoro 2014). Building on extant research and theory, we formulated and tested Hypothesis 3.

Hypothesis 3: *Given options A and B, a normative isomorphic pressure towards A increases the probability that public employees will choose A, even if A is inferior to B.*

Focusing more specifically on the rich stream of studies on isomorphism in public organizations, which mirrors the broader discussions in general management literature, we found that it has been concerned predominantly with organizational convergence and has focused on the organizational characteristics that are conducive to homogenization (D'Aunno, Sutton, and Price 1991; Frumkin and Galaskiewicz 2004; Bovaird and Downe 2006; Ashworth et al. 2007; Andrews 2011; Fink 2011; Fay and Zavattaro 2016). Along similar lines, when scholars have looked at individuals within public organizations, they have typically done so by identifying and testing their specific attributes, such as the structural embeddedness of top political officials (Villadsen 2011) or their professional socialization (Teodoro 2014).

Moreover, this rich stream of research also often warns us that isomorphic pressures may not necessarily be conducive to solutions that are technically superior. The intellectual endeavor of our study is to nail down empirically whether and how this is the case and, consistently with the micro-foundations perspective, to do so by focusing on the level of individual decisions rather than on the individual attributes of civil servants. Our research design allowed us to embrace this challenge and, in what follows, we explain the rationale behind the choice to conduct eight survey experiments and the qualitative inquiry, and describe the stages of this research journey.

3.2 Methods

Participants, Design, and Procedures

To test our hypotheses, we conducted eight randomized controlled trials on three independent samples of Italian public sector workers recruited through *Qualtrics*. Sample sizes were 204 (Experiment 1), 396 (Experiment 2), and 164 (Experiments 3a, 3 b, 4a, 4 b, 5a and 5b). Randomized experiments have been described as “the most efficient tool that researchers and program evaluators have at their disposal to obtain an unbiased estimate of the average effect caused by an intervention of some kind” (Bellé and Cantarelli 2017:3). The experimental part of the study was supplemented by the analysis of interview data collected from participants in the third sample. Indeed, although randomized controlled trials are well-suited for testing causal relations between two variables (i.e., molar causation), they do not necessarily help illuminate the chain reaction linking causes to their effects (i.e., molecular causation) (Shadish, Cook, and Campbell 2002). In light of this inherent limitation, we supplemented the experimental phase of our project with a qualitative enquiry in an attempt to gain a deeper understanding of the motives driving subjects’ decisions (Appendix B reports the English translation of the experimental scenarios).

The Experiments

The eight randomized control trials shared a common design. In each experiment, subjects were randomly assigned to one of four scenarios (control, coercive isomorphism, mimetic isomorphism, and normative isomorphism) and asked to decide between two options. The control scenario provided only information about differences in performance between the two options. In addition to this information, the three isomorphic scenarios prompted participants that one of the two options

was preferred by either a higher authority (coercive), their best-performing peers (mimetic), or their professional networks (normative).

Experiment 1

Subjects were asked to imagine themselves as the superintendent of a school district who had to choose between two management software packages. Subjects in the control were informed that both software packages had been approved by the Italian Ministry of Education, had the same price, and that there was no evidence of one being better than the other. In other words, we primed subjects to think that the two software packages were equivalent. In addition to the same information provided to the control group, subjects in the three isomorphic conditions were informed that one of the two options was: suggested by the guidelines issued by the Ministry of Education (coercive), going to be adopted by the school districts with the highest reputation nationwide (mimetic), or recommended by the professional association of superintendents to which they belonged (normative).

Experiment 2

Experiment 2 replicated Experiment 1 with only one variation: whereas subjects in Experiment 1 were primed to believe that the two software packages were equivalent in all respects, subjects in Experiment 2 were informed that one's performance was slightly worse than the other's performance. We held constant across the two experiments all other features and procedures described earlier.

Experiments 3a and 3b

Experiments 3a and 3b were variations of Experiment 2 on a different sample and in different decision settings. In particular, subjects in Experiment 3a were not prompted to imagine themselves as superintendents, but instead were asked to

suggest that their own institutions adopt one of two managerial software packages. Everything else in Experiment 3a was the same as in Experiment 2. Participants in Experiment 3b had to choose between two training programs. Subjects in the control condition of Experiment 3b read that (1) the two training programs provided the same number of credits, had the same schedule of classes, and required the same effort; and (2) the performance assessments issued by a reliable independent agency indicated that one training program was slightly worse than the other. In addition to the same information provided to respondents in the control group, subjects in the three isomorphic conditions read that the inferior training program was suggested by their Human Resources Director (coercive), their best colleagues (mimetic), and their former colleagues/classmates (normative).

Experiments 4a and 4b

Experiments 4a and 4b replicated Experiments 3a and 3b, respectively, with the only exception that we specified the performance dimensions along which one option was slightly worse than the other. We listed such dimensions in the text. More precisely, in addition to the same information that respondents read in Experiment 3a, subjects in Experiment 4a were informed that one software package was inferior to the other in terms of ease of use, speed, accuracy, and technical support. Respondents in Experiment 4b were told that the inferior training program performed slightly worse in terms of practical usefulness, quality of the instructors, quality of the content, and efficacy of the teaching methodologies.

Experiments 5a and 5b

Unlike in the other experiments, in Experiments 5a and 5b participants were not explicitly informed about differences in performance between the two options. Instead, subjects were presented with a table reporting performance scores for

each of the two options along the same dimensions indicated in Experiment 4. Therefore, whereas public workers who participated in the previous experiments were primed about performance differences, subjects in Experiment 5 had to infer this information themselves, based on numeric data reported in a table format.

We situated Experiments 1 and 2 in the context of school district management because this setting lends itself to an investigation of all three isomorphic pressures. First, superintendents have an asymmetric relationship with a centralized organization - i.e., the Ministry of Education - which exerts power upon and provides resources to the schools, and the Ministry typically provides recommendations in the form of guidelines to school districts in Italy (coercive). Second, school district superintendents are the target of high-powered performance management and benchmarking practices (e.g., league tables), which have the potential to trigger reputation concerns and imitative behaviors (mimetic). Third, superintendents are characterized by a high degree of executive professionalism as they are routinely engaged in professional associations (normative). Furthermore, investigating how isomorphic pressures play out in the context of a school district may be relevant on its own, given that education is the largest industry within the Italian public sector and certainly a prominent one in most countries. In Experiment 1, the two software packages were described as performing equally well; the target software was portrayed as inferior in Experiment 2. Therefore, Experiment 2 complements Experiment 1 by providing a more robust test of isomorphism. Experiments 3 through 5 were designed to test the external validity of our findings for different decision settings (i.e., choice between two managerial software packages and two training programs) and different ways of presenting information about the inferiority of the target option (i.e., generic statement, performance dimensions listed in a textual format, performance scores shown in a table). Overall, the rationale behind our manipulation of sub-optimality was twofold. First, we opted for a low-intensity

treatment to ensure greater contextual realism. Given the features of our scenarios, a wider performance gap between the superior and inferior options would have been quite unrealistic. The second reason for choosing a low-intensity treatment was to induce variation in decisions across subjects, which is a prerequisite to conduct meaningful statistical tests.

The Qualitative Inquiry

The experimental part of Experiments 3 through 5 was followed by a set of open-ended questions. These allowed exploring how participants interpret the role and features of the institutions or benchmark groups to operationalize the different isomorphic pressures in the experiment, as well as to provide a more in-depth and nuanced account of isomorphic pressures at work. In particular, questions invited respondents to describe how they conceived, both in general and with specific reference to decision making, the roles of ministerial guidelines and professional associations as well as that of their best-performing peers or the public organizations with the highest reputation nationwide. Second and relatedly, questions invited respondents to describe whether and why they would follow the advice of these same institutions or groups (i.e., ministerial guidelines, colleagues or agencies with the best reputation and professional associations) in case they recommended the inferior solution between two alternatives, either software packages or training programs.

The transcripts' analysis combined deductive a priori broad themes emerging from the experimental phase, such as the types of isomorphism, with data-driven inductive coding from our questions (Krippendorff 2004). Coding was not geared towards measuring rates of responses or frequency of specific constructs, but rather at allowing original themes to emerge directly from the transcripts (Fereday and Muir-Cochrane 2006). It was performed with the support of the software program

Figure 8: First-order and second-order codes emerging from open-ended questions.



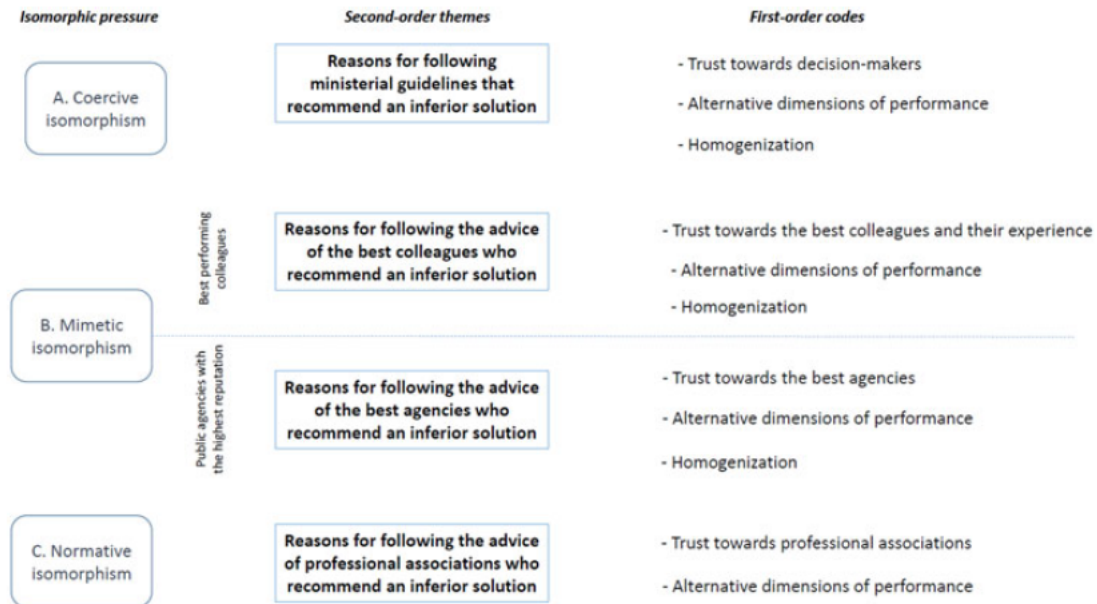
ATLAS.ti. The final data structure is provided in Figures 8 and 9. Additional illustrative quotations are provided in the Appendices C and D.

3.3 Results

Experiments

Table 12 lists the demographic characteristics for the three samples of public workers who participated in the study. Within sample one and sample two, participants in the four experimental conditions (i.e., control, coercive, mimetic, and normative) did not differ in terms of average age, proportion of females, proportion of managers (i.e., subjects who managed at least one subordinate), distribution by public sector industry, and distribution by type of degree. Within sample three, a series of chi-square tests and t-tests unveiled some significant differences across experimental arms in terms of managerial status, industry of employment, and age.

Figure 9: First-order and second-order codes emerging from open-ended questions about ex-post accounts of the reasons for following a specific pressure.



In particular, the proportion of managers in the mimetic arm of Experiment 4a was 34 percentage points lower than in the coercive arm ($p = .001$) and 33 percentage points lower than in the normative arm ($p = .002$). Also, the distribution of subjects by industry of employment varied among conditions in Experiment 4a ($Pearsonchi2(9) = 19.57, p = 0.021$). Lastly, the average age of respondents in the coercive arm of Experiment 4b was 5.07 years lower than in the mimetic arm ($p = .024$) and 4.51 years lower than in the normative arm ($p = .044$).

Figure 10 displays the percentage of subjects choosing the software package that was encouraged for the three treated groups, but not for the control group, in Experiment 1. The percentages of public workers opting for the option were as follows: 43.6% in the control group (i.e., in the absence of any isomorphic pressures); 85.7% in the coercive condition (i.e., when the encouraged software was suggested by guidelines from the Ministry of Education); 85.0% in the mimetic condition (i.e.,

Table 12: Demographic Characteristics of Respondents, by Sample

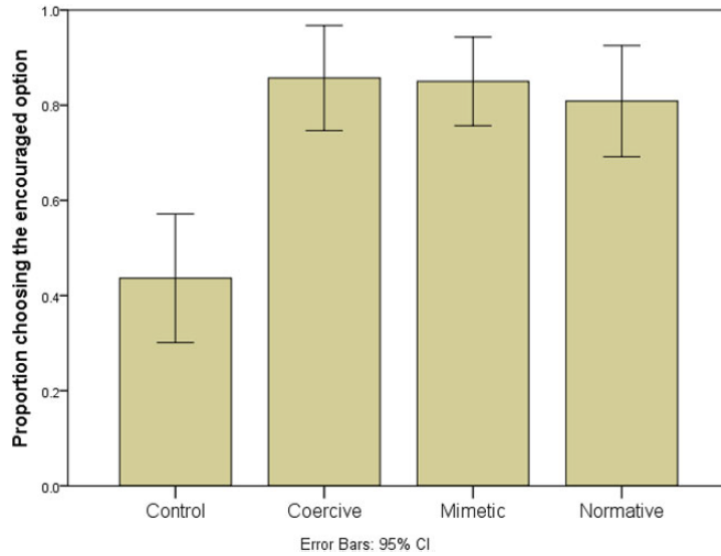
	Sample One: Experiment 1	Sample Two: Experiment 2	Sample Three: Experiments 3a, 3b, 4a, 4b, 5a, 5b
<i>N</i>	204	396	164
Age: $\mu(\sigma)$ in years	45.9 (10.0)	43.4 (10.2)	49.0 (9.98)
Female %	68.6	48.7	50
Manager %	43.1	74.8	50.6
Public sector industry			
Healthcare %	16.3	16.4	17.1
Education %	67.5	39.1	43.9
General administration %	7.4	28.3	25.6
Other %	8.9	16.2	13.4
Degree %	79.4	73.7	60.3
Scientific %	38.2	37.1	26.8
Humanities %	41.2	36.6	33.5

when the software was adopted by the school district with the highest reputation); and 80.9% in the normative condition (i.e., when the software was recommended by the superintendents' association).

The results of a logistic regression showed that each of the three isomorphic pressures that we manipulated significantly increased the odds that a public employee would choose the encouraged software package. Compared to the control group, the odds of choosing the encouraged option increased by 7.75 times ($p < .001$) under a coercive pressure, by 7.32 times ($p < .001$) under a mimetic pressure, and by 5.45 times ($p < .001$) under a normative pressure (Table 13). The effects of all three experimentally induced isomorphic pressures on the probability of choosing the encouraged software were significantly different from zero and statistically indistinguishable from one another in size.

Figure 11 shows the results of Experiment 2, where participants had to choose between two software packages, one of which was slightly inferior to the other. The percentages of subjects choosing the inferior software package for each of the four experimental groups were as follows: 11.9% in the control group, 41.3% in the

Figure 10: Proportion of subjects choosing the encouraged software, by isomorphic pressure (Experiment 1). Note: The two software packages are equivalent.

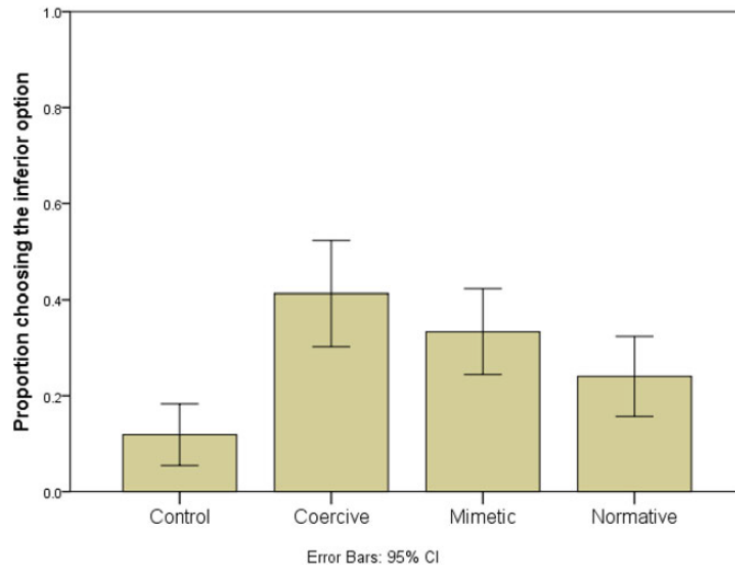


coercive condition, 33.3% in the mimetic condition, and 24.0% in the normative condition.

The results of a logistic regression showed that each of the three isomorphic manipulations significantly increased the odds that a public employee would choose the slightly inferior software. Compared to the control group, the odds of choosing the worse option increased by 5.21 times ($p < .001$) under a coercive pressure, by 3.71 times ($p < .001$) under a mimetic pressure, and by 2.35 times ($p < .05$) under a normative pressure (Table 13). A series of Wald tests indicated that the probability of choosing the inferior option was higher under a coercive rather than a normative pressure ($p < .05$). We did not find any other significant differences among the three types of pressures.

The left column of Figure 12 shows the results of Experiment 3a. A logistic regression showed that, relative to the control condition, the odds that participants would suggest that their own organizations adopt the inferior software package were

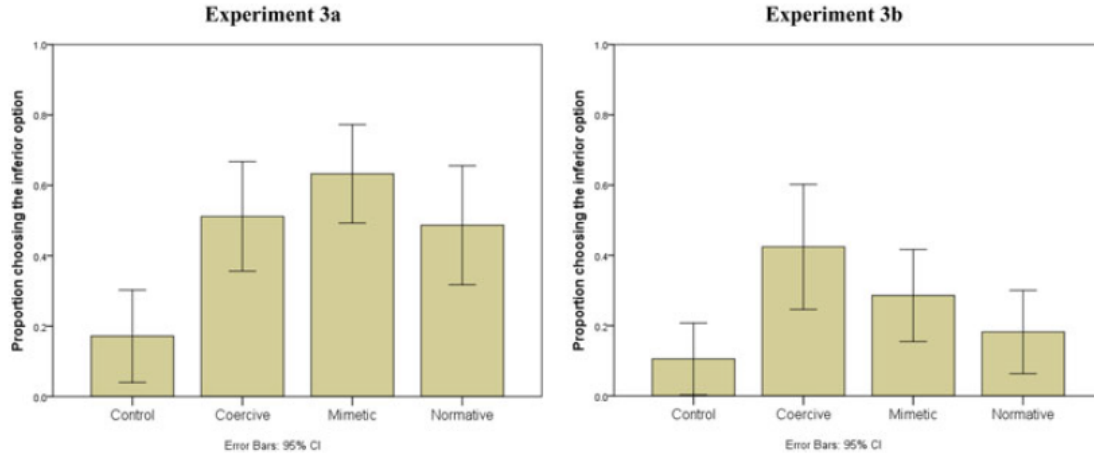
Figure 11: Proportion of subjects choosing the inferior software, by isomorphic pressure (Experiment 2). Note: One software is inferior to the other - generic statement.



5.06 times higher under a coercive pressure ($p = .003$), 8.32 times higher under a mimetic pressure ($p < .001$), and 4.58 times higher under a normative pressure ($p = .006$) (Table 13).

The right column of Figure 12 plots the findings from Experiment 3b, in which subjects had to indicate which one of two training programs they would attend, by isomorphic condition. The pattern of results for Experiment 3b is consistent with what we observed in the previous studies, with the only exception being the normative condition. More precisely, as in the previous randomized trials, the coercive and mimetic isomorphic pressures had the expected effects of increasing the likelihood that respondents would select the inferior option as compared to the control condition. Participants in the normative condition were no more likely to choose the worse training program than their counterparts in the control condition ($p = .333$). The results of a logistic regression revealed that, relative to the control group, the odds of choosing the inferior training program were 6.26 times higher

Figure 12: Proportion of subjects choosing the inferior option, by isomorphic pressure.



for participants who were told that their Human Resources directors had suggested the inferior option (coercive) ($p = .004$) and 3.40 times higher for subjects informed that their best colleagues had decided to attend the inferior program (mimetic) ($p = .047$) (Table 13).

Figure 13 reports the proportions of subjects who selected the inferior software package for each of the four manipulations in Experiment 4a. Unlike in Experiment 3a, where participants were told that the performance of one software was slightly worse than the performance of the other software, in Experiment 4a subjects were informed that one software package was inferior to the other in terms of ease of use, speed, accuracy, and technical support.

Estimated coefficients from a logistic regression indicated that, compared to the control group, the odds of opting for the inferior software were 3.85 higher for public employees assigned to the coercive scenario ($p = .018$) and 3.12 higher for subjects in the mimetic pressure ($p = .053$). The sign of the coefficient for the normative group was positive but the associated p-value was .223 (Table 13). Therefore, unlike in Experiment 3a, we did not find support for Hypothesis 3.

Figure 13: Proportion of subjects choosing the inferior software, by isomorphic pressure.

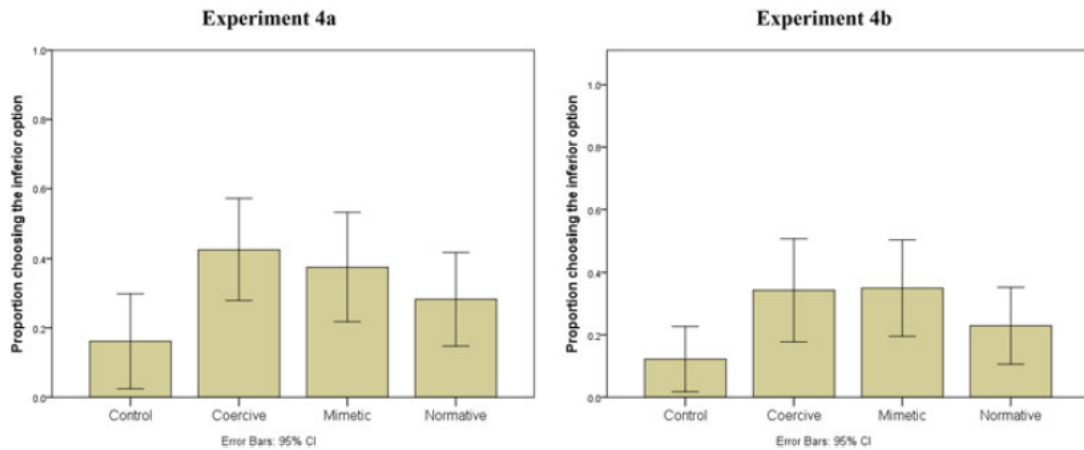
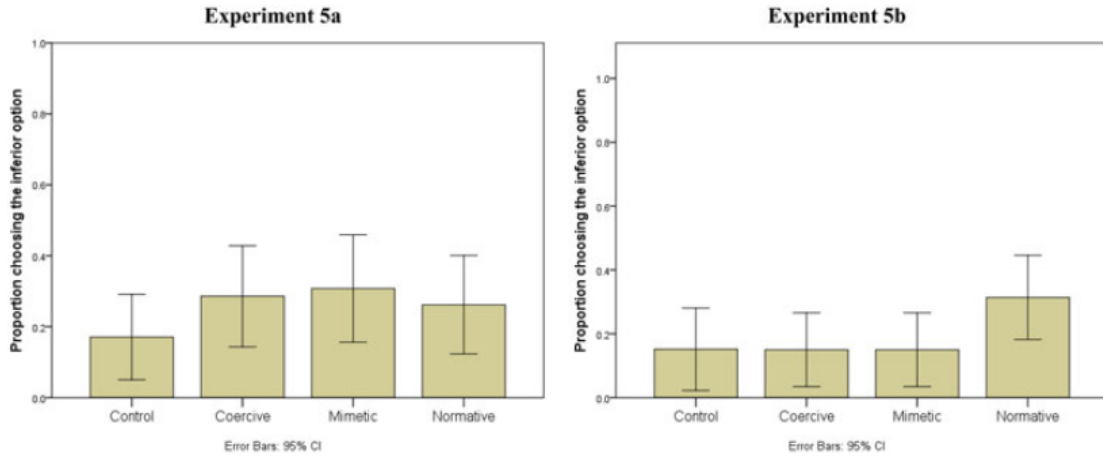


Figure 13 also shows the findings from Experiment 4b in which public employees were told that an independent agency had rated one training program as slightly worse than the other along four performance dimensions (i.e., practical usefulness, quality of the instructors, quality of the content, and efficacy of the teaching methodologies). The pattern of results in Experiment 4b was the same as in Experiments 3b and 4a. The coercive manipulation increased the odds that subjects would prefer the inferior training program by 3.76 ($p = .026$). The odds of choosing the inferior option were 3.88 times higher for participants exposed to the mimetic scenario relative to their peers in the control condition ($p = .020$). Again, the coefficient associated with the normative pressure had the expected sign but did not reach significance at the conventional level ($p = .196$) (Table 13).

Whereas in all of the previous experiments subjects were presented with textual information about the relative performance of the options, participants in Experiment 5 were shown a table reporting numerical scores of the two alternatives for the same performance dimensions listed in Experiment 4. Figure 14 suggests that none of the isomorphic manipulations had a significant on the odds of choosing the

Figure 14: Proportion of subjects choosing the inferior software, by isomorphic pressure.



inferior software package (Table 13).

The lack of any significant impact from the isomorphic pressures is even clearer in the right column of Figure 14, which displays the percentages of subjects opting for the inferior training program by condition. Results of the logistic regression underlying Experiment 5b are reported in Table 13.

Table 13 reports the findings of a series of meta-analyses that we conducted to synthesize results across our randomized trials 2–5. Separately, for each of the three isomorphic pressures, we meta-analyzed the effect sizes from our eight experiments. Overall, the odds of choosing the target option went up by 3.64 times under a coercive pressure, relative to the control condition ($p < .001$). We observed similar meta-analytic results for mimetic isomorphism ($OR = 3.33, p < .001$) and normative isomorphism ($OR = 2.37, p < .001$). The results of a meta-analysis that combined effects across the three isomorphic pressures indicated that the overall impact of isomorphism was 3.06 ($p < .001$).

Table 13: Changes in the Odds (w.r.t. Control) of Choosing the Encouraged Inferior Option, by Type of Isomorphic Pressure and by Experiment (2 through 5)

		OR	z	p
Coercive	Experiment 2	5.21	4.32	0.000
	Experiment 3a	5.06	2.99	0.003
	Experiment 4a	3.85	2.36	0.018
	Experiment 5a	1.94	1.24	0.217
	Experiment 3b	6.26	2.89	0.004
	Experiment 4b	3.76	2.22	0.026
	Experiment 5b	0.99	-0.02	0.985
	Coercive Overall	3.64	6.40	0.000
Mimetic	Experiment 2	3.71	3.57	0.000
	Experiment 3a	8.32	3.94	0.000
	Experiment 4a	3.12	1.94	0.053
	Experiment 5a	2.16	1.42	0.155
	Experiment 3b	3.40	1.99	0.047
	Experiment 4b	3.88	2.33	0.020
	Experiment 5b	0.99	-0.02	0.985
	Mimetic Overall	3.33	6.01	0.000
Normative	Experiment 2	2.35	2.22	0.025
	Experiment 3a	4.58	2.74	0.006
	Experiment 4a	2.05	1.22	0.223
	Experiment 5a	1.72	1.00	0.317
	Experiment 3b	1.89	0.98	0.333
	Experiment 4b	2.14	1.29	0.196
	Experiment 5b	2.56	1.64	0.100
	Normative Overall	2.37	4.28	0.000
	Isomorphism	3.06	9.63	0.000
	Overall			

Note: Odds ratios (OR) estimated using fixed effects.

Qualitative Inquiry

Transcript analysis of the answers that our informants provided to the open questions allowed us to identify recurring themes that helped explain how individual expectations about ministerial guidelines, the advice of highly rated colleagues or agencies, as well as the advice of professional associations lead to isomorphic responses. Recurring themes included individual expectations about the rigor and usefulness of decisions enacted as authoritative guidelines, expectations about the support for decision making, expectations about intrinsic qualities of the benchmark groups (best colleagues or best agencies) and, last, the perception that homogenization was important for civil servants.

The presentation of our most significant findings is organized by type of isomorphic pressures (Figure 8), while also offering a systematic collection of additional evidence supporting the coding in Appendix C.

Coercive Isomorphism

Respondents' opinions of ministerial guidelines, with specific reference to decision making, revolve around three main elements. First, guidelines are trusted for their expected intrinsic qualities that represent the *attributes of decision*. They are *based on reliable studies* "carried out in depth and that no single public agency could afford." The second source of legitimacy for ministerial guidelines lies in the expertise of their authors. Guidelines are in fact *drafted by experts*, by personnel "with technical know-how." Moreover, they are *based on practical experience* and, in particular, they "perform the function of identifying services and standards that have been tested already and that can be therefore considered reliable." It may be argued that, for some of our respondents, ministerial guidelines are received uncritically and considered legitimate, even when in contrast with objective inputs to decision making, such as those presented in our experimental scenarios.

Second, guidelines offer *support for the decision-making process*. They are a *reference point*, by providing certainty and by presenting useful terms of reference that can easily be retrieved and consulted. The decision process is facilitated. In the words of one of our informants: “Guidelines are important because they make life easier for whoever is called to make a decision.” Respondents also pointed to the importance of guidelines as an *enabler of professional conduct*, following which civil servants are kept on track, are safe in terms of “avoiding being sued for nothing,” and have their professional profile delimited thanks to the “boundaries into what may otherwise be too vague as the behavioral expectations towards civil servants.”

Third, our analysis revealed that the *homogenization* of practice enabled by the guidelines is seen as a value per se. In particular, they *guarantee uniformity in the public sector*. According to our informants, guidelines are meant to harmonize the activities of the public sector and “are crucial to avoid each agency operating according to its own way and to avoid useless duplication in looking for the same information.”

Mimetic Isomorphism

Consistent with the way in which we operationalized mimetic isomorphism in our experiment, in the open-ended questions we explored the expectations for the recommendations from the best colleagues and the best agencies, especially as far as decisions were concerned. The two sets of findings are substantially coherent and revolve around three main elements.

First, respondents seem to trust the intrinsic attributes of decisions made by the best colleagues or the by highly rated agencies. Their expectations are that the “best colleagues,” who are taught to invest time and effort into their choices, will make decisions *based on serious and committed assessment*: “Typically best performing

colleagues display high levels of effort in choosing,” and their decisions are “based on sound evaluation criteria and perform well in their decisions (...) result(ing) from a serious and thoughtful evaluation.” Again, this seems to silence the critical thinking of some respondents. Similarly, respondents trust the *recommendations of the agencies with the highest reputation*, as they invest time and effort to produce a *careful assessment* of costs, benefits and the reliability of the provider. Moreover, their decisions are *based on objective criteria*, such as “efficiency and merit, and you can always trust these principles,” which can somehow “be considered universal criteria.” A variation on this theme is related to the expectation that, albeit technically inferior, the solutions adopted by the best colleagues are *based also on practical applicability*; i.e., they “choose based on the utility of a specific option, such as a training program or a software.” In a similar vein, the recommendations of the best agencies are expected to be based on a fit with the needs: “They choose on the basis of the needs of personnel and on the updates that are needed to keep up with progress and changes.”

Second, the recommendations of the best-performing colleagues and of the public agencies with the best reputation support the decision-making process. The best-performing colleagues do so by giving advice to those with scarce experience or skills. The agencies with the highest reputations do so by specifying the criteria for selection: “If you look at the best agencies and follow them, you can’t be wrong. At the end of the day they have built their reputation with commitment and sacrifice and they select on the basis of their know-how that is then shared.”

Third, our analysis shows that respondents align their decisions with those of the best colleagues and agencies based on their consideration for the intrinsic qualities of these peers and institutions. Therefore, they are willing to follow the decisions made by the “role models,” by those who are capable. The best colleagues are also the pioneers, those able to anticipate trends and, as reported by one of our

respondents, “I always think that the best colleagues understand earlier and better the validity of a specific training program or a specific software, and I am therefore inclined to follow them.” Moving from the best colleagues to the best agencies, decisions made by pioneering agencies offer valuable guidance because “the best agencies are those that implement innovation first and reform processes and that display efficiency.” In a similar vein, “the best agencies are those you can trust. On the basis of trust, you follow what they do.”

Normative Isomorphism

Respondents conform to the recommendations of professional associations, for they promote *decisions aimed at developing the competences* of their members: “Professional associations inform, train, guarantee and offer guidelines to their members, and they provide suggestions that are in line with this purpose.” Interestingly, our informants repeatedly pointed to the importance of following *decisions aimed at protecting members’ rights*, which is in line with what is considered perhaps the most important mandate of a professional association. In the vivid words of one of our respondents, “The main role of professional associations is to defend, yes, I mean defend, the interests of their members. I don’t know if they always choose or recommend the best in absolute terms, but certainly they choose based on the protection of their members.” Informants highlighted the importance of professional associations for offering support to decision making, again mentioning that this is based on technical expertise and referring to the fact that “the decisions and recommendations of professional associations are safe” for their members. Last, professional associations guarantee a homogenization among civil servants with effects that are relevant within and outside the category: “Following the recommendations of professional associations helps maintaining a certain internal conduct within the category and this will also be recognized outside.”

Isomorphic Pressures Leading to the Choice of an Inferior Solution

We now turn to the analysis of the answers offered by our respondents to the questions of whether and why they would follow ministerial guidelines, colleagues or agencies with the best reputation, and professional associations, even when they recommended the inferior solution between two alternatives (i.e., software packages or training programs). We posit that the responses to these explicit questions enrich our analysis of isomorphic pressures by illustrating, in the respondents' own words, their ex-post account of the reasons for following a specific pressure. While the data structure (Figure 9) and a systematic collection of evidence supporting the coding (Appendix D) present the ex-post justification per each type of isomorphic pressure, we summarize in the following the three themes that emerged from the analysis.

One theme is *trust towards the institution or group recommending the solution*. The type of explanations included in this theme appears quite tautological, as effectively illustrated by this answer: "I trust what the ministry recommends, no matter what." Explanations along this line rest on the claim that confidence and belief towards the ministry issuing guidelines, colleagues and agencies, or professional associations automatically result in the acceptance of their advice. We have identified as a second emerging theme the *alternative dimensions of performance* elicited by our respondents. Examples of this theme point to the possibility that the apparent inferiority of the solution covers advantages for the public agency or the civil servants: "Perhaps they recommend an inferior training or software because, while being slightly inferior, it better meets the needs of the agency." As put by one of our informants: "An example shall clarify my view on why these recommendations should be followed, irrespectively of the performance of a specific product. If I need a motorcycle and, at the same price, they offer me a scooter or a big motorbike, the latter may certainly be better and enable me to

do more things, but if at the end of the day I only need to travel around in the city center, I will end up choosing the scooter (...).” The third emerging theme is *compatibility with existing standards*, whereby the choice of the best solution may lead to a misalignment with “the great majority of colleagues (and) would end up being useless or even counterproductive.”

3.4 Discussion and conclusions

By adopting a micro-foundations perspective (Powell and Colyvas 2008; Bitektine and Haack 2015; Felin et al. 2015; Grimmelikhuijsen et al. 2017), the current study has brought individuals, their choices, and their interaction with collective actors into an analysis of isomorphic pressures. In so doing, it complements previous findings predominantly focused on the characteristics of organizations (D’Aunno et al. 1991; Frumkin and Galaskiewicz 2004; Ashworth et al. 2007; Fink 2011; Fay and Zavattaro 2016) and on the dynamics or mechanisms through which isomorphic pressures unfold (Radaelli 2000; Lodge and Wegrich 2005; Currie and Suhomlinova 2006; Gong and Xiao 2017). It also shows what individuals actually do, rather than inferring which of their characteristics are more conducive to isomorphism (Villadsen 2011; Teodoro 2014).

Our research design, which patterned these premises, enabled the appreciation of the interplay between individuals and collective actors, such as government agencies and professional organizations acting upon some socialized legitimacy judgments (Bitektine and Haack 2015). In particular, the study has revealed that mechanisms connecting individual judgments to those collective actors, such as trust in the recommending institution or group, speculation about alternative performance dimensions, and compatibility with existing standards, are the main drivers of suboptimal decision making.

Scholars have expressed doubt that isomorphism will lead to performance im-

provements in public organizations (Radaelli 2000; Pollitt 2001; Frumkin and Galaskiewicz 2004; Ashworth et al. 2007; Kallio and Kuoppakangas 2013; Ammons and Roenigk 2015). Our findings provide empirical ammunition to substantiate this claim unequivocally. In fact, studies on the convergence of phenomena that ranged from arrangements such as privatization to practices such as university branding suffered from the lack of a clear baseline and had a more general difficulty in measuring outcomes. In contrast, the simplified scenario of our randomized trial allowed us to determine, less contentiously, that isomorphic pressures may lead to inferior solutions.

As previously mentioned, it has been recognized that the three isomorphic pressures are not easy to disentangle empirically since, although each of them involves a separate mechanism, they may be working together and have simultaneous effects (Mizruchi and Fein 1999). Very recently, for example, a qualitative study by Gong and Xiao (2017), set in local governments in China, identified empirically the three different isomorphic pressures. Our survey experiments allowed us to make an incremental step in the direction of isolating the three effects, because subjects in each of the treated groups were exposed to just one type of isomorphic pressure. Consistent with the majority of experiments of this type, the choice of operations requires judgment calls and a certain degree of discretion (e.g., Harrison and List 2004). With specific regard to isomorphic pressures, operations can be formulated in different ways depending on the context, the organizations, and the actors involved. Even inside the same context, more than one valid operation is often available.

Our research design seems to be well-equipped to meet internal validity requirements and, therefore, to establish a causal link between our isomorphic pressures' constructs and the propensity to make suboptimal decisions. However, our results should be interpreted in light of some limitations that pave the way for future

research. First, our experimental manipulations of the three types of pressure are inherently incomparable because there is no single scale against which the dosages of our interventions can be measured (Shadish et al. 2002). As a consequence, the observed effects cannot be univocally attributed to either the type of pressure or the magnitude of the manipulation. We were partially able to overcome this limitation by supplementing our randomized controlled trials with qualitative work. Second, our design is prone to external validity threats that are common to most experimental work (Baekgaard et al. 2015). For instance, the use of an abstract framing, an imposed set of rules, and a pool of participants who self-selected themselves for the study detracts from the generalizability of the findings beyond the study setting to more naturally occurring environments. External validity threats may be mitigated by the fact that subjects are real public sector workers. Furthermore, our design allowed exploiting variation in participants' industry of employment within government. Nonetheless, replications of our experiments might test our findings in different contexts and with different operationalizations of isomorphic pressures. As for the coercive pressure, for example, given that in our case it refers to influence from sources of authority that have potentially coercive power, it would be interesting to test for different operationalizations varying in the degree of direct coercion involved in the pressure (e.g., authoritative mandates or directives with sanctions attached, as contrasted with recommendations). In this respect, additional variations of the source of performance information and the way in which this is released to decision makers might be very informative.

The qualitative findings help to refine our analysis by accounting for the reasons why public employees follow a specific pressure. Cross-cutting themes among the three types of isomorphism include, first, the attributes of decisions; i.e., the intrinsic qualities that our respondents assign to ministerial guidelines as well as to the recommendations of the peers or agencies with the highest reputation. Not

only is their reliance on “objective studies” considered a plus, but also their fit with organizational needs. In a similar vein, a second theme is the support for the decision process, which highlights the importance that respondents attribute to the availability of devices, such as guidelines and recommendations, which identify criteria for their choice and may compensate for any lack of skills or experience. Qualities of and support for the decision also characterize the role of recommendations provided by professional associations, which we employed to operationalize normative isomorphism. Interestingly, however, these specific recommendations are conceived more as a shield for public employees, as an instrument that may guarantee their rights, rather than uniquely increase their performance or enable their professional development. Our findings also point to the importance of what we have called “homogenization,” which has emerged as a theme of both coercive and normative isomorphism. This peculiar theme, confirmed in the ex-post account of the reasons for following a specific pressure, refers to the awareness of public employees that both behavioral convergence and compatibility among technical solutions in government are positive results per se, even if they come at the price of selecting an inferior solution. We posit that such awareness could be further investigated, as it casts the choice of an inferior solution as the result of thorough considerations rather than of contingent impulses.

To conclude, in line with previous research, our results suggest that isomorphic pressures may overpower evidence-based arguments indicating unambiguously superior solutions. However, it also provides a nuanced account of the underlying reasons. This should not be taken as an invitation to surrender decisions in public organizations to uncertainty and vagaries (Fischer 1990; Sanderson 2002). Instead, a major implication we can draw is the importance of employing institutional devices that trigger isomorphic pressures to channel and reinforce evidence-based content. In other words, we should not assume that evidence will speak for itself (Majone

1989; Mele, Compagni, and Cavazza 2014). Our results highlight the importance of relying on devices such as professional associations, epistemic communities, and orchestrated benchmarking exercises to reinforce the likelihood of adopting technically superior decisions, as well as to foresee and handle preconceptions that may derail such adoptions.

4 Conclusions

A robust discussion has begun on decision-making biases in public management, administration, and policy. Most scholars have investigated how citizens make informed assessments of government policies (see, e.g., Andersen and Hjortskov 2015; Geys and Sørensen 2017; Grosso, Charbonneau, and Van Ryzin 2017; Jilke, Van Ryzin, and Van de Walle 2016; Marvel 2016; Olsen 2017). Fewer studies, however, have reviewed the decision processes of public managers (e.g., Bellé, Cantarelli, and Belardinelli 2017) and policy makers (e.g., Moynihan and Lavertu 2012).

Nonetheless, while exploring the systematic errors of citizens' judgments about government services is important, it is equally relevant and timely to understand whether and how cognitive biases affect public workers responsible for designing, managing, and implementing public policies and managerial procedures. This thesis has provided with new empirical evidence on these dynamics.

Public managers and employees who participated in the randomized trials included in chapter 1 were found to be prone to a number of cognitive biases like framing, anchoring, proportion dominance, status quo, and asymmetric dominance. In particular, the framing of outcomes influenced decisions across policy and management domains. In addition, public employees were prone to an anchoring bias when setting standards for responsiveness, were more likely to stick to a suboptimal status quo as the number of superior alternatives increased, and tended to put more effort into activities that affected higher percentages of beneficiaries, even if the absolute number of affected clients was constant. Lastly, decisions changed when a decoy was present, proving an asymmetric dominance effect. In chapter 2, performance information use by local public managers was investigated. They were more likely to be subject to framing effects under ex post uses of performance information (e.g., service evaluation) than ex ante (e.g., resource allocation). Inter-

estingly, asking them to justify for their choices did not work as a debiasing strategy in their ex post decisions. Finally, chapter 3 investigates the micro-foundations of isomorphic pressures which lead to suboptimal decision-making in the public sphere. The eight experiments included in this last chapter showed that isomorphic pressures indeed can lead to inferior solutions. However, in this setting, giving in to isomorphic pressures may actually be interpreted as a rational strategy, for example to cope with uncertainty. The qualitative findings helped identify public managers and employees who gave in to isomorphic pressures because they thought that was the best they could do with the available information and those who actually fell under bandwagoning and therefore inconsistently revealed their preferences.

Future research in public administration might further adopt behavioral lenses to deepen our knowledge about the effects of biased public decision-making on our societies. A recent systematic review by Battaglio and colleagues (2019) identifies two main reasons to elevate behavioral public administration. On the one hand, behavioral science may help generate research that has the potential to advance theory and inspire practice at once within the field of public administration. On the other hand, behavioral public administration may facilitate the dialogue between the field of public administration and other disciplines in which the use of behavioral science has long been well established. “In the marketplace of ideas, public administration suffers from an unfavorable balance because we tend to import way more from other fields than we export. Strengthening behavioral public administration may help narrow this trade gap and enhance our contribution to broader social science” (Battaglio et al. 2019, pp. 315-316).

References

- Abadie, Alberto, and Sebastien Gay. 2006. The Impact of Presumed Consent Legislation on Cadaveric Organ Donation: A Cross-Country Study. *Journal of Health Economics* 25(4): 599–620.
- Abrahamson, Eric. 1991. Managerial fads and fashions: The diffusion and rejection of innovations. *Academy of Management Review* 16(3): 586–612.
- Alevy, Jonathan E., Craig E. Landry, and John A. List. 2015. Field Experiments on the Anchoring of Economic Valuations. *Economic Inquiry* 53(3): 1522–38.
- Ammons, David N., and William C. Rivenbark. 2008. Factors Influencing the Use of Performance Data to Improve Municipal Services: Evidence from the North Carolina Benchmarking Project. *Public Administration Review* 68(2): 304–18.
- Ammons, David N., and Dale J. Roenigk. 2014. Benchmarking and interorganizational learning in local government. *Journal of Public Administration Research and Theory* 25(1): 309–335.
- Andersen, Simon Calmar, and Morten Hjortskov. 2015. Cognitive biases in performance evaluations. *Journal of Public Administration Research and Theory* 26(4): 647–662.
- Andrews, Matt. 2011. Which organizational attributes are amenable to external reform? An empirical study of African public financial management. *International Public Management Journal* 14(2): 131–156.
- Ariely, Dan. 2010. *Predictably irrational, revised and expanded edition*. Harper Collings Publishers, London.
- Ariely, Dan, and Thomas S. Wallsten. 1995. Seeking Subjective Dominance in Multidimensional space: An Explanation of the Asymmetric Dominance Effect. *Organizational Behavior and Human Decision Processes* 63(3): 223–32.
- Artinger, Florian, Malte Petersen, Gerd Gigerenzer, and Jürgen Weibler. 2015. Heuristics as Adaptive Decision Strategies in Management. *Supplement 1, Journal of Organizational Behavior* 36: S33–52.
- Ashworth, Rachel, George Boyne, and Rick Delbridge. 2009. Escape from the iron cage? Organizational change and isomorphic pressures in the public sector. *Journal of Public Administration Research and Theory* 19(1): 165–187.
- Baekgaard, Martin. 2017. Prospect theory and public service outcomes: Examining risk preferences in relation to public sector reforms. *Public Administration* 95(4): 927–942.
- Baekgaard, Martin, Julian Christensen, Casper Mondrup Dahlmann, Asbjørn Mathiasen, and Niels Bjørn Grund Petersen. 2017. The role of evidence in politics: Motivated reasoning and persuasion among politicians. *British Journal of Political Science*, 1–24.

Baekgaard, Martin, and Søren Serritzlew. 2016. Interpreting performance information: Motivated reasoning or unbiased comprehension. *Public Administration Review* 76(1): 73-82.

Baekgaard, Martin, et al. 2015. Conducting experiments in public management research: A practical guide. *International public management journal* 18(2): 323-342.

Baiman, Stanley. 1982. Agency Research in Managerial Accounting: A Survey. *Journal of Accounting Literature* 1: 154-213.

Baron, Jonathan. 1997. Confusion of Relative and Absolute Risk in Valuation. *Journal of Risk and Uncertainty* 14(3): 301-9.

Barrows, Samuel, Michael Henderson, Paul E. Peterson, and Martin R. West. 2016. Relative performance information and perceptions of public service quality: Evidence from American school districts. *Journal of Public Administration Research and Theory* 26(3): 571-583.

Barrows, Samuel, Margaret Henderson, Paul E. Peterson, and Martin R. West. 2016. Relative Performance Information and Perceptions of Public Service Quality: Evidence from American School Districts. *Journal of Public Administration Research and Theory* 26(3): 571-83.

Bartels, Daniel M. 2006. Proportion Dominance: The Generality and Variability of Favoring Relative Savings over Absolute Savings. *Organizational Behavior and Human Decision Processes* 100(1): 76-95.

Battaglio Jr, R. Paul, Paolo Belardinelli, Nicola Bellé, and Paola Cantarelli. 2019. Behavioral Public Administration ad fontes: A Synthesis of Research on Bounded Rationality, Cognitive Biases, and Nudging in Public Organizations. *Public Administration Review* 79(3): 304-320.

Beattie, Jane, Jonathan Baron, John C. Hershey, and Mark D. Spranca. 1994. Psychological Determinants of Decision Attitude. *Journal of Behavioral Decision Making* 7(2): 129-44.

Behn, Robert D. 2003. Why Measure Performance? Different Purposes Require Different Measures. *Public Administration Review* 63(5): 586-606.

Bellé, Nicola. 2010. Così fan tutte? Adoption and rejection of performance-related pay in Italian municipalities: A cross-sector test of isomorphism. *Review of Public Personnel Administration* 30(2): 166-188.

Bellé, Nicola, and Paola Cantarelli. 2017. What Causes Unethical Behavior? A Meta-Analysis to Set an Agenda for Public Administration Research. *Public Administration Review* 77(3): 327-339.

Bellé, Nicola, Paola Cantarelli, and Paolo Belardinelli. 2017. Cognitive Biases in Performance Appraisal: Experimental Evidence on Anchoring and Halo Effects With Public Sector Managers and Employees. *Review of Public Personnel Administration* 37(3): 275-294.

ben-Aaron, James, et al. 2017. Transparency by conformity: A field experiment evaluating openness in local governments. *Public Administration Review* 77(1): 68-77.

Bennett, Mark W. 2014. Confronting Cognitive “Anchoring Effect” and “Blind Spot” Biases in Federal Sentencing: A Modest Solution for Reforming a Fundamental Flaw. *Journal of Criminal Law and Criminology* 104(3): 489–534.

Bernoulli, Daniel. 1954. Exposition of a new theory on the measurement of risk. *Econometrica* 22: 23–36. (Original work published 1738).

Bitektine, Alex, and Patrick Haack. 2015. The ‘macro’ and the ‘micro’ of legitimacy: Toward a multilevel theory of the legitimacy process. *Academy of Management Review* 40(1): 49-75.

Blumenthal-Barby, Jennifer S., and Heather Krieger. 2015. Cognitive biases and heuristics in medical decision making: a critical review using a systematic search strategy. *Medical Decision Making* 35(4): 539-557.

Borgida, Eugene, and Beth Howard-Pitney. 1983. Personal Involvement and the Robustness of Perceptual Salience Effects. *Journal of Personality and Social Psychology* 45(3): 560–70.

Borins, Sandford. 2000. Loose cannons and rule breakers, or enterprising leaders? Some evidence about innovative public managers. *Public Administration Review* 60(6): 498-507.

Bouckaert, Geert. 1993. Efficiency Measurement from a Management Perspective: A Case of the Civil Registry Office in Flanders. *International Review of Administrative Sciences* 59(1): 11–27.

Bouwman, Robin, and Stephan Grimmelikhuijsen. 2016. Experimental Public Administration from 1992 to 2014: A Systematic Literature Review and Ways Forward. *International Journal of Public Sector Management* 29(2): 110–31.

Bovaird, Tony, and James Downe. 2006. N generations of reform in UK local government: compliance and resistance to institutional pressures. *International Public Management Journal* 9(4): 429-455.

Butler, Daniel M., and David E. Broockman. 2011. Do Politicians Racially Discriminate against Constituents? A Field Experiment on State Legislators. *American Journal of Political Science* 55(3): 463–77.

Cabinet Office, Behavioural Insights Team. 2014. *EAST: Four Simple Ways to Apply Behavioural Insights*. http://38r8om2xjhh125mw24492dir.wpengine.netdna-cdn.com/wp-content/uploads/2015/07/BIT-Publication-EAST_FA_WEB.pdf [Last Accessed July 8, 2018].

Caplin, Andrew, and Paul W. Glimcher. 2014. Basic methods from neoclassical economics. In *Neuroeconomics*, edited by Paul W. Glimcher and Ernst Feher, 3-17. London, UK: Elsevier.

Carroll, Gabriel D., James J. Choi, David Laibson, Brigitte C. Madrian, and

Andrew Metrick. 2009. Optimal Defaults and Active Decisions. *Quarterly Journal of Economics* 124(4): 1639–74.

Carter, Craig R., Lutz Kaufmann, and Alex Michel. 2007. Behavioral Supply Management: A Taxonomy of Judgment and Decision-Making Biases. *International Journal of Physical Distribution Logistics Management* 37(8): 631–69.

Chaiken, Shelly. 1980. Heuristic versus Systematic Information Processing and the Use of Source versus Message Cues in Persuasion. *Journal of Personality and Social Psychology* 39(5): 752–66.

Chen, Zhe, and Simon Kemp. 2015. Anchoring Effects in Simulated Academic Promotion Decisions: How the Promotion Criterion Affects Ratings and the Decision to Support an Application. *Journal of Behavioral Decision Making* 28(2): 137–48.

Compagni, Amelia, Valentina Mele, and Davide Ravasi. 2015. How early implementations influence later adoptions of innovation: Social positioning and skill reproduction in the diffusion of robotic surgery. *Academy of Management Journal* 58(1): 242-278.

Cornelissen, Joep P., and Mirjam D. Werner. 2014. Putting Framing in Perspective: A Review of Framing and Frame Analysis across the Management and Organizational Literature. *Academy of Management Annals* 8(1): 181–235.

Cronqvist, Henrik, and Richard H. Thaler. 2004. Design Choices in Privatized Social-Security Systems: Learning from the Swedish Experience. *American Economic Review* 94(2): 424–28.

Currie, Graeme, and Olga Suhomlinova. 2006. The impact of institutional forces upon knowledge sharing in the UK NHS: the triumph of professional power and the inconsistency of policy. *Public Administration* 84(1): 1-30.

D'Aunno, Thomas, Robert I. Sutton, and Richard H. Price. 1991. Isomorphism and external support in conflicting institutional environments: A study of drug abuse treatment units. *Academy of management journal* 34(3): 636-661.

Dacin, M. Tina. 1997. Isomorphism in context: The power and prescription of institutional norms. *Academy of management journal* 40(1): 46-81.

Davis, Mark A., and Philip Bobko. 1986. Contextual Effects on Escalation Processes in Public Sector Decision Making. *Organizational Behavior and Human Decision Processes* 37(1): 121–38.

Deephouse, David L. 1996. Does isomorphism legitimate?. *Academy of management journal* 39(4): 1024-1039.

Demski, Joel S., and Gerald A. Feltham. 1976. *Cost Determination: A Conceptual Approach*. Ames: Iowa State University Press.

DiMaggio, Paul. 2012. *Constructing an organizational field as a professional project: The case of US art museums*. Institutional theory in organizationa studies. Sage Publications.

DiMaggio, Paul J. and Walter. W. Powell. 1983. The iron cage revisited: Insti-

tutional isomorphism and collective rationality in organizational fields. *American Sociological Review* 48(2): 147-160.

Ditillo, Angelo, Mariannunziata Liguori, Mariafrancesca Sicilia, and Ilena Steccolini. 2015. Control Patterns in Contracting-Out Relationships: It Matters What You Do, Not Who You Are. *Public Administration* 93(1): 212–29.

Duflo, Esther, William Gale, Jeffrey Liebman, Peter Orszag, and Emmanuel Saez. 2006. Saving Incentives for Low- and Middle-Income Families: Evidence from a Field Experiment with HR Block. *Quarterly Journal of Economics* 121(4): 1311–46.

Dunleavy, Patrick. 1996. *Democracy, Bureaucracy and Public Choice*. New York: Harvester Wheatsheaf.

Englich, Birte, Thomas Mussweiler, and Fritz Strack. 2006. Playing Dice with Criminal Sentences: The Influence of Irrelevant Anchors on Experts' Judicial Decision Making. *Personality and Social Psychology Bulletin* 32(2): 188–200.

Erlandsson, Arvid, Fredrik Björklund, and Martin Bäckström. 2015. Emotional Reactions, Perceived Impact and Perceived Responsibility Mediate the Identifiable Victim Effect, Proportion Dominance Effect and In-Group Effect Respectively. *Organizational Behavior and Human Decision Processes* 127: 1–14.

Fagley, Nancy S., and Paul M. Miller. 1991. The Effects of Framing, Problem Variations, and Providing Rationale on Choice. *Personality and Social Psychology Bulletin* 17(5): 517–22.

Fay, Daniel L., and Staci M. Zavattaro. 2016. Branding and isomorphism: the case of Higher Education. *Public Administration Review* 76(5): 805-815.

Felin, Teppo, Nicolai J. Foss, and Robert E. Ployhart. 2015. The microfoundations movement in strategy and organization theory. *The Academy of Management Annals* 9(1): 575-632.

Fereday, Jennifer, and Eimear Muir-Cochrane. 2006. Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International journal of qualitative methods* 5(1): 80-92.

Festinger, Leon. 1957. *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.

Fetherstonhaugh, David, Paul Slovic, Stephen Johnson, and James Friedrich. 1997. Insensitivity to the Value of Human Life: A Study of Psychophysical Numbing. *Journal of Risk and Uncertainty* 14(3): 283–300.

Fink, Simon. 2011. A contagious concept: Explaining the spread of privatization in the telecommunications sector. *Governance* 24(1): 111-139.

Fischer, Frank. 1990. *Technocracy and the Politics of Expertise*. Newbury Park, CA: Sage Publications.

Fligstein, Neil. 1990. *The transformation of corporate control*. Cambridge, MA: Harvard University Press.

Fligstein, Neil. 1991. The structural transformation of American industry: An institutional account of the causes of diversification in the largest firms, 1919-1979. In W.W. Powell and P. J. DiMaggio (eds), *The new institutionalism in organizational analysis*. Chicago, IL: University of Chicago Press, pp. 294-310.

Frumkin, Peter, and Joseph Galaskiewicz. 2004. Institutional isomorphism and public sector organizations. *Journal of Public Administration Research and Theory* 14(3): 283-307.

Furnham, Adrian, and Hua Chu Boo. 2011. A Literature Review of the Anchoring Effect. *Journal of Socio-Economics* 40(1): 35-42.

George, Bert, Martin Baekgaard, Adeliën Decramer, Mieke Audenaert, and Stijn Goeminne. 2018. Institutional Isomorphism, Negativity Bias and Performance Information Use by Politicians: A Survey Experiment. *Public Administration*. Published electronically on February 7. <https://doi.org/10.1111/padm.12390>.

George, Bert, Sebastian Desmidt, Poul A. Nielsen, and Martin Baekgaard. 2016. Rational Planning and Politicians' Preferences for Spending and Reform: Replication and Extension of a Survey Experiment. *Public Management Review*. 19(9): 1251-71.

Geys, Benny, and Rune J. Sørensen. 2017. Never change a winning policy? Public sector performance and politicians' preferences for reforms. *Public Administration Review* 78(2): 206-216.

Gigerenzer, Gerd, Peter M. Todd, and the ABC Research Group. 1999. *Simple heuristics that make us smart*. Oxford University Press.

Gimeno, Javier, et al. 2005. Explaining the clustering of international expansion moves: A critical test in the US telecommunications industry. *Academy of Management Journal* 48(2): 297-319.

Glynn, Mary Ann, and Rikki Abzug. 2002. Institutionalizing identity: Symbolic isomorphism and organizational names. *Academy of Management journal* 45(1): 267-280.

Gong, Ting, and Hanyu Xiao. 2017. The Formation and Impact of Isomorphic Pressures: Extravagant Position-Related Consumption in China. *Governance* 30(3): 387-405.

Gordon, R. A. Y., Martin Kornberger, and Stewart R. Clegg. 2009. Power, rationality and legitimacy in public organizations. *Public Administration* 87(1): 15-34.

Greenwood, Royston, Roy Suddaby, and Christopher R. Hinings. 2002. Theorizing change: The role of professional associations in the transformation of institutionalized fields. *Academy of management journal* 45(1): 58-80.

Grimmelikhuijsen, Stephan, Sebastian Jilke, Asmus L. Olsen, and Lars Tummers. 2017. Behavioral public administration: Combining insights from public administration and psychology. *Public Administration Review* 77(1): 45-56.

- Grosso, Ashley, Étienne Charbonneau, and Gregg G. Van Ryzin. 2017. How Citizens Respond to Outputs, Outcomes, and Costs: A Survey Experiment About an HIV/AIDS Program. *International Public Management Journal* 20(1): 160-181.
- Guthrie, Chris, Jeffrey J. Rachlinski, and Andrew J. Wistrich. 2001. Inside the Judicial Mind. *Cornell Law Review* 86(4): 777-830.
- Haas, Peter M. 1992. Introduction: epistemic communities and international policy coordination. *International organization* 46(1): 1-35.
- Harkness, Allan R., Kenneth G. DeBono, and Eugene Borgida. 1985. Personal Involvement and Strategies for Making Contingency Judgements: A Stake in the Dating Game Makes a Difference. *Journal of Personality and Social Psychology* 49(1): 22-32.
- Harrison, Glenn W., and John A. List. 2004. Field Experiments. *Journal of Economic Literature* 42(4): 1009-55.
- Hartley, Jean, and James Downe. 2007. The shining lights? Public service awards as an approach to service improvement. *Public Administration* 85(2): 329-353.
- Haunschild, Pamela R., and Anne S. Miner. 1997. Modes of interorganizational imitation: The effects of outcome salience and uncertainty. *Administrative Science Quarterly*, 472-500.
- Haveman, Heather A. 1993. Follow the leader: Mimetic isomorphism and entry into new markets. *Administrative science quarterly*: 593-627.
- Henri, Jean-François. 2006. Organizational Culture and Performance Measurement Systems. *Accounting, Organizations and Society* 31(1): 77-103.
- Highhouse, Scott. 1996. Context-Dependent Selection: The Effects of Decoy and Phantom Job Candidates. *Organizational Behavior and Human Decision Processes* 65(1): 68-76.
- Hood, Christopher. 1991. *A Public Management for All Seasons?* *Public Administration* 69(1): 3-19.
- Huber, Joel, John W. Payne, and Christopher Puto. 1982. Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis. *Journal of Consumer Research* 9(1): 90-98.
- Huber, Joel, and Christopher Puto. 1983. Market Boundaries and Product Choice: Illustrating Attraction and Substitution Effects. *Journal of Consumer Research* 10(1): 31-44.
- Jacobs, Rick, and Steve W. J. Kozlowski. 1985. A Closer Look at Halo Error in Performance Ratings. *Academy of Management Journal* 28(1): 201-12.
- Jilke, Sebastian, Gregg G. Van Ryzin, and Steven Van de Walle. 2016. Responses to decline in marketized public services: An experimental evaluation of choice overload. *Journal of Public Administration Research and Theory* 26(3): 421-432.
- John, Peter, and Gerry Stoker. 2019. Rethinking the role of experts and expertise in behavioural public policy. *Policy Politics* 47(2): 209-225.

- Johnson, Eric J., and Daniel Goldstein. 2003. Do Defaults Save Lives? *Science* 302: 1338–39.
- Julnes, Patria de Lancer, and Marc Holzer. 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, Daniel. 2011. *Thinking, fast and slow*. Macmillan.
- . 2002. Nobel Prize Lecture. *Maps of bounded rationality: A perspective on intuitive judgment and choice*.
- . 2000. A Psychological Point of View: Violations of Rational Rules as a Diagnostic of Mental Processes. *Behavioral and Brain Sciences* 23(5): 681–83.
- Kahneman, Daniel, and Gary Klein. 2009. Conditions for Intuitive Expertise: A Failure to Disagree. *American Psychologist* 64(6): 515–26.
- Kahneman, Daniel, and Amos Tversky. 1979. Prospect Theory: An Analysis of Decision Under Risk, *Econometrica* 47: 263-291.
- Kallio, Tomi J., and Päivikki Kuoppakangas. 2013. Bandwagoning municipal enterprises: institutional isomorphism and the search for the Third Way. *Policy Studies* 34(1) 19-35.
- Katz, Richard. 2017. “Interviewing a Nobel Prize Winner.” Econvue, October 10. <https://econvue.com/pulse/interviewing-nobel-prize-winner> [accessed May 19, 2018].
- Kelman, Steven, Ronald Sanders, and Gayatri Pandit. 2016. “I Won’t Back Down?” Complexity and Courage in Government Executive Decision Making. *Public Administration Review* 76(3): 465-471.
- Kieser, Alfred. 1997. Rhetoric and myth in management fashion. *Organization* 4(1): 49-74.
- Koriat, Asher, Sarah Lichtenstein, and Baruch Fischhoff. 1980. Reasons for Confidence. *Journal of Experimental Psychology: Human Learning Memory* 6(2): 107–18.
- Krippendorff, Klaus. 2018. *Content analysis: An introduction to its methodology*. Sage publications.
- Kroll, Alexander. 2015. Drivers of Performance Information Use: Systematic Literature Review and Directions for Future Research. *Public Performance and Management Review* 38(3): 459–486.
- . 2014. Why performance information use varies among public managers: Testing manager-related explanations. *International Public Management Journal* 17(2): 174-201.
- Kühberger, Anton. 1998. The Influence of Framing on Risky Decisions: A Meta-Analysis. *Organizational Behavior and Human Decision Processes* 75(1): 23–55.

Lægreid, Per, Paul G. Roness, and Kristin Rubecksen. 2007. Modern management tools in state agencies: The case of Norway. *International Public Management Journal* 10(4): 387-413.

Larrick, Richard P. 2004. Debiasing. In *Blackwell Handbook of Judgment and Decision Making*, edited by Derek J. Koehler and Nigel Harvey, 316-38. Malden, MA: Blackwell.

Lavertu, Stéphane. 2016. We All Need Help: “Big Data” and the Mismeasure of Public Administration. *Public Administration Review* 76(6): 864–72.

LeBoeuf, Robyn A., and Eldar Shafir. 2003. Deep Thoughts and Shallow Frames: On the Susceptibility to Framing Effects. *Journal of Behavioral Decision Making* 16(2): 77–92.

Levin, Irwin P., and Daniel P. Chapman. 1990. Risk Taking, Frame of Reference, and Characterization of Victim Groups in AIDS Treatment Decisions. *Journal of Experimental Social Psychology* 26(5): 421–34.

Levin, Irwin P., Sandra L. Schneider, and Gary J. Gaeth. 1998. All Frames Are Not Created Equal: A Typology and Critical Analysis of Framing Effects. *Organizational Behavior and Human Decision Processes* 76(2): 149–88.

Levitt, Steven D., and John A. List. 2007. What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World? *Journal of Economic Perspectives* 21(2): 153–74.

Liguori, Mariannunziata, Mariafrancesca Sicilia, and Ileana Steccolini. 2012. Some Like It Non-financial . . . Politicians’ and Managers’ Views on the Importance of Performance Information. *Public Management Review* 14(7): 903–22.

Linde, Jona, and Barbara Vis. 2017. Do Politicians Take Risks Like the Rest of Us? An Experimental Test of Prospect Theory under MPs. *Political Psychology* 38(1): 101–17.

Lodge, Martin, and Kai Wegrich. 2005. Control over government: Institutional isomorphism and governance dynamics in German public administration. *Policy Studies Journal* 33(2): 213-233.

Majone, Giandomenico. 1989. *Evidence, argument, and persuasion in the policy process*. Yale University Press.

Marvel, John D. 2016. Unconscious Bias in Citizens’ Evaluations of Public Sector Performance. *Journal of Public Administration Research and Theory* 26(1): 143–58.

—. 2015. Public Opinion and Public Sector Performance: Are Individuals’ Beliefs About Performance Evidence-Based or the Product of Anti-Public Sector Bias? *International Public Management Journal* 18(2): 209–27.

Meekings, Alan. 1995. Unlocking the Potential of Performance Measurement: A Guide to Practical Implementation. *Public Money Management* 15(4): 5–12.

Meier, Kenneth J., Søren C. Winter, Laurence J. O’Toole, Nathan Favero, and

- Simon C. Andersen. 2015. The validity of subjective performance measures: School principals in Texas and Denmark. *Public administration* 93(4): 1084-1101.
- Mele, Valentina, Amelia Compagni, and Marianna Cavazza. 2013. Governing through evidence: a study of technological innovation in health care. *Journal of Public Administration Research and Theory* 24(4): 843-877.
- Melkers, Julia, and Katherine Willoughby. 2005. Models of Performance Measurement Use in Local Governments: Understanding Budgeting, Communication, and Lasting Effects. *Public Administration Review* 65(2): 180-90.
- Meyer, John W., and Brian Rowan. 1977. Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*: 340-363.
- Meyer, John, W. Richard Scott, and David Strang. 1987. Centralization, fragmentation, and school district complexity. *Administrative Science Quarterly*: 186-201.
- Mizruchi, Mark S., and Lisa C. Fein. 1999. The social construction of organizational knowledge: A study of the uses of coercive, mimetic, and normative isomorphism. *Administrative science quarterly* 44(4): 653-683.
- Moynihan, Donald P. 2009. The Network Governance of Crisis Response: Case Studies of Incident Command Systems. *Journal of Public Administration Research and Theory* 19(4): 895-915.
- Moynihan, Donald P., Sergio Fernandez, Soonhee Kim, Kelly M. LeRoux, Suzanne J. Piotrowski, Bradley E. Wright, and Kaifeng Yang. 2011. Performance Regimes amidst Governance Complexity. *Supplement 1, Journal of Public Administration Research and Theory* 21: 141-55.
- Moynihan, Donald, Pamela Herd, and Hope Harvey. 2014. Administrative burden: Learning, psychological, and compliance costs in citizen-state interactions. *Journal of Public Administration Research and Theory* 25(1): 43-69.
- Moynihan, Donald P., and Stéphane Lavertu. 2012. Cognitive biases in governing: Technology preferences in election administration. *Public administration review* 72(1): 68-77.
- Moynihan, Donald P., and Sanjay K. Pandey. 2010. The Big Question for Performance Management: Why Do Managers Use Performance Information? *Journal of Public Administration Research and Theory* 20(4): 849-66.
- Moynihan, Donald P., Sanjay K. Pandey, and Bradley E. Wright. 2012. Prosocial Values and Performance Management Theory: Linking Perceived Social Impact and Performance Information Use. *Governance* 25(3): 463-83.
- Nielsen, Poul A., and Martin Baekgaard. 2013. 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.
- Nielsen, Poul A., and Donald P. Moynihan. 2017. How Do Politicians Attribute

- Bureaucratic Responsibility for Performance? Negativity Bias and Interest Group Advocacy. *Journal of Public Administration Research and Theory* 27(2): 269-283.
- . 2016. Romanticizing Bureaucratic Leadership? The Politics of How Elected Officials Attribute Responsibility for Performance. *Governance* 30(4): 541-59.
- Niskanen, William A. 1971. *Bureaucracy and Representative Government*. New Brunswick, NJ: Transaction Publishers.
- O'Donnell, Ed, and Joseph J. Schultz, Jr. 2005. The Halo Effect in Business Risk Audits: Can Strategic Risk Assessment Bias Auditor Judgment about Accounting Details? *Accounting Review* 80(3): 921-39.
- Olsen, Asmus L. 2015. Citizen (Dis)Satisfaction: An Equivalence Framing Study. *Public Administration Review* 75(3): 469-478.
- Olsen, Asmus. 2017. Human interest or hard numbers? Experiments on citizens' selection, exposure, and recall of performance information. *Public Administration Review* 77(3): 408-420.
- Organization for Economic Cooperation and Development (OECD). (2017), *Behavioural Insights and Public Policy: Lessons from Around the World*. OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264270480-en> [Last accessed July 8, 2018]
- Orr, Dan, and Chris Guthrie. 2006. Anchoring, Information, Expertise, and Negotiation: New Insights from Meta-Analysis. *Ohio State Journal on Dispute Resolution* 21(3): 597-628.
- Orru, Marco, Nicole W. Biggart, and Gary G. Hamilton. 1991. *Organizational Isomorphism in East Asia: Broadening the New Institutionalism*. Institute of Governmental Affairs, University of California, Davis.
- Osborne David, and Ted Gaebler. 1992. *Reinventing Government*. New York: Penguin.
- Palmer, Donald A., P. Devereaux Jennings, and Xueguang Zhou. 1993. Late adoption of the multidivisional form by large US corporations: Institutional, political, and economic accounts. *Administrative Science Quarterly*: 100-131.
- Petty, Richard E., John T. Cacioppo, and Rachel Goldman. 1981. Personal Involvement as a Determinant of Argument-Based Persuasion. *Journal of Personality and Social Psychology* 41(5): 847-55.
- Pina, Vicente, Lourdes Torres, and Sonia Royo. 2010. Is e-government promoting convergence towards more accountable local governments?. *International Public Management Journal* 13(4): 350-380.
- Pitts, David W., et al. 2010. What drives the implementation of diversity management programs? Evidence from public organizations. *Journal of Public Administration Research and Theory* 20(4): 867-886.
- Pollitt, Christopher. 2001. Convergence: the useful myth?. *Public administration* 79(4): 933-947.

- Pollitt, Christopher, and Geert Bouckaert. 2011. *Public Management Reform: A Comparative Analysis—New Public Management, Governance, and the Neo-Weberian State*. 3rd ed. Oxford: Oxford University Press.
- Powell, W. W., and Colyvas, J. A. 2008. *Microfoundations of institutional theory*. The Sage handbook of organizational institutionalism. SAGE (276-298).
- Rabin, Matthew. 1998. Psychology and Economics. *Journal of Economic Literature* 36(1): 11–46.
- Rachlinski, Jeffrey J., Chris Guthrie, and Andrew J. Wistrich. 2006. Inside the Bankruptcy Judge’s Mind. *Boston University Law Review* 86(5): 1227–65.
- Radaelli, Claudio M. 2000. Policy transfer in the European Union: institutional isomorphism as a source of legitimacy. *Governance* 13(1): 25-43.
- Redelmeier, Donald A., and Eldar Shafir. 1995. Medical Decision Making in Situations That Offer Multiple Alternatives. *Journal of the American Medical Association* 273(4): 302–305.
- Sanderson, Ian. 2002. Evaluation, policy learning and evidence-based policy making. *Public administration* 80(1): 1-22.
- Sapoznik, Gustavo, Redelmeier Donald, Ruff Christian, and Philippe N. Tobler. 2016. Cognitive biases associated with medical decisions: a systematic review. *BMC medical informatics and decision making* 16(1): 138.
- Schedler, Kuno. 2003. “. . . And Politics?” Public Managements in the Light of Two Rationalities. *Public Management Review* 5(4): 533–50.
- Schooler, Jonathan W., and Joseph Melcher. 1995. The Ineffability of Insight. In *The Creative Cognition Approach*, edited by Steven M. Smith, Thomas B. Ward, and Ronald A. Finke, 97–133. Cambridge, MA: MIT Press.
- Schooler, Jonathan W., Stellan Ohlsson, and Kevin Brooks. 1993. Thoughts beyond Words: When Language Overshadows Insight. *Journal of Experimental Psychology: General* 122(2): 166–83.
- Scott, Peter. 1995. *The meanings of mass higher education*. McGraw-Hill Education (UK).
- Sedikides, Constantine, Dan Ariely, and Nils Olsen. 1999. Contextual and Procedural Determinants of Partner Selection: Of Asymmetric Dominance and Prominence. *Social Cognition* 17(2): 118–139.
- Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton, Mifflin and Company.
- Showers, Carolin, and Nancy Cantor. 1985. Social Cognition: A Look at Motivated Strategies. *Annual Review of Psychology* 36: 275–305.
- Sieck, Winston, and J. Frank Yates. 1997. Exposition Effects on Decision Making: Choice and Confidence in Choice. *Organizational Behavior and Human Decision Processes* 70(3): 207–19.

- Simmons, Joseph P., Robyn A. LeBoeuf, and Leif D. Nelson. 2010. The Effect of Accuracy Motivation on Anchoring and Adjustment: Do People Adjust from Provided Anchors? *Journal of Personality and Social Psychology* 99(6): 917–32.
- Simon, Andrew F., Nancy S. Fagley, and Jennifer G. Halleran. 2004. Decision Framing: Moderating Effects of Individual Differences and Cognitive Processing. *Journal of Behavioral Decision Making* 17(2): 77–93.
- Simon, Herbert A. 1978. Nobel Prize Lecture. Rational Decision-Making in Business Organizations. http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/1978/simonlecture.html [Last accessed July 8, 2018].
- . 1956. Rational Choice and The Structure of the Environment. *Psychological Review* 63(2): 129–38.
- . 1955. A Behavioral Model of Rational Choice. *Quarterly Journal of Economics* 69(1): 99–118.
- Simon, Herbert A. 1954. *Centralization vs. Decentralization in Organizing the Controllers' Department*. New York: Controllershship Foundation Inc.
- . 1947. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*. New York: Macmillan.
- Simons, Robert. 1990. The Role of Management Control Systems in Creating Competitive Advantage: New Perspectives. *Accounting, Organizations and Society* 15(1–2): 127–43.
- Slaughter, Jerel E., Jessica Bagger, and Andrew Li. 2006. Context Effects on Group-Based Employee Selection Decisions. *Organizational Behavior and Human Decision Processes* 100(1): 47–59.
- Slaughter, Jerel E., Evan F. Sinar, and Scott Highhouse. 1999. Decoy Effects and Attribute-Level Inferences. *Journal of Applied Psychology* 84(5): 823–28.
- Slovic, Paul, Melissa L. Finucane, Ellen Peters, and Donald G. MacGregor. 2007. The Affect Heuristic. *European Journal of Operational Research* 177(3): 1333–52.
- Smith, Stephen M., and Irwin P. Levin. 1996. Need for Cognition and Choice Framing Effects. *Journal of Behavioral Decision Making* 9(4): 283–90.
- Smith, Vernon L. 2003. Constructivist and ecological rationality in economics. *The American Economic Review* 93(3): 465–508.
- Speklé, Roland F., and Frank H. Verbeeten. 2014. The Use of Performance Measurement Systems in the Public Sector: Effects on Performance. *Management Accounting Research* 25(2): 131–46.
- Sprinkle, Geoffrey B. 2003. Perspectives on Experimental Research in Managerial Accounting. *Accounting, Organizations and Society* 28(2–3): 287–318.
- Stanovich, Keith E. and Richard F. West. 2000. Individual Differences in Reasoning: Implications for the Rationality Debate? *Behavioral and Brain Sciences* 23(5): 645–65.
- Stinessen, Leif. 1985. The Influence of Verbalization on Problem Solving.

Scandinavian Journal of Psychology 26(1): 342–47.

Sunstein, Cass R., and Richard H. Thaler. Libertarian paternalism is not an oxymoron.” *The University of Chicago Law Review* (2003): 1159-1202.

Takemura, Kazuhisa. 1994. Influence of Elaboration on the Framing of Decisions. *Journal of Psychology* 128(1): 33–39.

—. 1993. The Effect of Decision Frame and Decision Justification on Risky Choice. *Japanese Psychological Research* 35(1): 36–40.

Taylor, Jeannette. 2011. Factors Influencing the Use of Performance Information for Decision Making in Australian State Agencies. *Public Administration* 89(4): 1316–34.

Teodoro, Manuel P. 2014. When professionals lead: Executive management, normative isomorphism, and policy implementation. *Journal of Public Administration Research and Theory* 24(4): 983-1004.

Thaler, Richard H. 2017. Nobel Prize Lecture. *From Cashews to Nudges: The Evolution of Behavioral Economics*. http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2017/thalerlecture.html [Last accessed July 8, 2018]

—. 2015. *Misbehaving: The Making of Behavioral Economics*. New York: W. W. Norton.

—. 1980. Toward a Positive Theory of Consumer Choice. *Journal of Economic Behavior and Organization* 1(1): 39–60.

Thaler, Richard H., and Shlomo Benartzi. 2001. Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.153.5567> [accessed May 19, 2018].

Thaler, Richard H. and Cass Sustein. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Heaven. Yale University Press.

Thorsteinson, Todd J., Jennifer Breier, Anna Atwell, Catherine Hamilton, and Monica Privette. 2008. Anchoring Effects on Performance Judgments. *Organizational Behavior and Human Decision Processes* 107(1): 29–40.

Tiessen, Peter, and John H. Waterhouse. 1983. Towards a Descriptive Theory of Management Accounting. *Accounting, Organizations and Society* 8(2–3): 251–67.

Tolbert, Pamela S., and Lynne G. Zucker. 1983. Institutional sources of change in the formal structure of organizations: The diffusion of civil service reform, 1880-1935. *Administrative Science Quarterly* 28(1): 22-39.

Tversky, Amos and Daniel Kahneman. 1986. Rational Choice and the Framing of Decisions. *Journal of Business* 59(4), pt. 2: S251–78.

—. 1981. The framing of decisions and the psychology of choice. *Science* 221(4481): 453-458.

—. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185(4157): 1124-1131.

- Tversky, Amos, and Eldar Shafir. 1992. Choice under Conflict: The Dynamics of Deferred Decision. *Psychological Science* 3(6): 358–61.
- Tversky, Amos, and Itamar Simonson. 1993. Context-Dependent Preferences. *Management Science* 39(10): 1179–89.
- Van Dooren, Wouter, Geert Bouckaert, and John Halligan. 2010. *Performance Management in the Public Sector*. New York: Routledge.
- Van Dooren, Wouter, and Steven Van de Walle, eds. 2008. *Performance Information in the Public Sector: How It Is Used*. Basingstoke: Palgrave Macmillan.
- Vandenbosch, Betty. 1999. An Empirical Analysis of the Association between the Use of Executive Support Systems and Perceived Organizational Competitiveness. *Accounting, Organizations and Society* 24(1): 77–92.
- Van Ryzin, Gregg G., Norma M. Riccucci, and Huafang Li. 2017. Representative Bureaucracy and Its Symbolic Effect on Citizens: A Conceptual Replication. *Public Management Review* 19(9): 1365–79.
- Vasudeva, Gurneeta. 2013. Weaving together the normative and regulative roles of government: How the Norwegian Sovereign Wealth Fund’s responsible conduct is shaping firms’ cross-border investments. *Organization Science* 24(6): 1662-1682.
- Villadsen, Anders R. 2011. Structural embeddedness of political top executives as explanation of policy isomorphism. *Journal of Public Administration Research and Theory* 21(4): 573-599.
- Vis, Barbara. 2011. Prospect Theory and Political Decision Making. *Political Studies Review* 9(3): 334–43.
- Vlaev, Ivo, Dominic King, Paul Dolan, and Ara Darzi. 2016. The theory and practice of “nudging”: changing health behaviors. *Public Administration Review* 76(4): 550-561.
- Walker, Richard M., Oliver James, and Gene A. Brewer. 2017. Replication, Experiments and Knowledge in Public Management Research. *Public Management Review* 19(9): 1221–34.
- Wang, Yi, Scott Highhouse, Christopher J. Lake, Nicole L. Petersen, and Thaddeus B. Rada. 2017. Meta-analytic investigations of the relation between intuition and analysis. *Journal of Behavioral Decision Making* 30(1): 15-25.
- Williamson, Ian O., and Daniel M. Cable. 2003. Organizational hiring patterns, interfirm network ties, and interorganizational imitation. *Academy of Management Journal* 46(3): 349-358.
- Wilson, James Q. 1989. *Bureaucracy: What Government Agencies Do and Why They Do It*. New York: Basic Books.
- Wilson, Timothy D., Christopher E. Houston, Kathryn M. Etling, and Nancy Brekke. 1996. A New Look at Anchoring Effects: Basic Anchoring and Its Antecedents. *Journal of Experimental Psychology: General* 125(4): 387–402.
- Wilson, Timothy D., and Jonathan W. Schooler. 1991. Thinking Too Much:

Introspection Can Reduce the Quality of Preferences and Decisions. *Journal of Personality and Social Psychology* 60(2): 181–92.

Wu, Chin-Shan, and Fei-Fei Cheng. 2011. The Joint Effect of Framing and Anchoring on Internet Buyers' Decision-Making. *Electronic Commerce Research and Applications* 10(3): 358–68.

Appendix A - Experimental scenarios, chapter 2

The text in italics displays the experimental manipulations. The text in italics in the square brackets is inserted instead of the corresponding italics text in the vignette.

Experiment Alpha

Imagine you are the general manager of the Sports and Education Department in an Italian municipality. *75% (75% – 95% scale) of sports facilities’ users in town is satisfied [25% (5% – 25% scale) of sports facilities’ users in town is dissatisfied]* with the service.

- Assess the performance of the sports facilities’ director by moving the cursor in the following scale, where 0 means a very bad performance, and 100 means an excellent performance.
- Assess the performance of the sports facilities by moving the cursor in the following scale, where 0 means a very bad performance, and 100 means an excellent performance.
- Indicate the effort you would dedicate to improving the performance of sports facilities by moving the cursor in the following scale, where 0 means the least effort, and 100 means the maximum effort.
- Indicate how you would change the amount of budget assigned to sports facilities the next year, by selecting one of the following options (from +50% to -50%).

Experiments Beta and Epsilon⁵

Imagine you are the general manager of the Sports and Education Department in an Italian municipality. Users of municipal sports facilities were asked to express their opinion about the service by choosing between two options: satisfied or dissatisfied. *75% (75% – 95% scale) of sports facilities’ users in town is satisfied [25% (5% – 25% scale) of sports facilities’ users in town is dissatisfied]* with the service.

- Assess the performance of the sports facilities’ director by moving the cursor in the following scale, where 0 means a very bad performance, and 100 means an excellent performance.

⁵The only difference between experiments Beta and Epsilon is that in the latter the order of decisions subjects have to make is random.

- Assess the performance of the sports facilities by moving the cursor in the following scale, where 0 means a very bad performance, and 100 means an excellent performance.
- Indicate the effort you would dedicate to improving the performance of sports facilities by moving the cursor in the following scale, where 0 means the least effort, and 100 means the maximum effort.
- Indicate how you would change the amount of budget assigned to sports facilities the next year, by selecting one of the following options (from +50% to -50%).

Experiment Gamma

Imagine you are the general manager of the Sports and Education Department in an Italian municipality. *75% (75% – 95% scale) of sports facilities’ users in town is satisfied [25% (5% – 25% scale) of sports facilities’ users in town is dissatisfied]* with the service.

- Assess the performance of the sports facilities’ director by moving the cursor in the following scale, where 0 means a very bad performance, and 100 means an excellent performance. Briefly justify your choice here.
- Assess the performance of the sports facilities by moving the cursor in the following scale, where 0 means a very bad performance, and 100 means an excellent performance. Briefly justify your choice here.
- Indicate the effort you would dedicate to improving the performance of sports facilities by moving the cursor in the following scale, where 0 means the least effort, and 100 means the maximum effort. Briefly justify your choice here.
- Indicate how you would change the amount of budget assigned to sports facilities the next year, by selecting one of the following options (from +50% to -50%). Briefly justify your choice here.

Experiment Delta

Imagine you are the general manager of the Sports and Education Department in an Italian municipality. Users of municipal sports facilities were asked to express their opinion about the service by choosing between two options: satisfied or dissatisfied. *75% (75% – 95% scale) of sports facilities’ users in town is satisfied [25% (5% – 25% scale) of sports facilities’ users in town is dissatisfied]* with the service.

- Assess the performance of the sports facilities' director by moving the cursor in the following scale, where 0 means a very bad performance, and 100 means an excellent performance. Briefly justify your choice here.
- Assess the performance of the sports facilities by moving the cursor in the following scale, where 0 means a very bad performance, and 100 means an excellent performance. Briefly justify your choice here.
- Indicate the effort you would dedicate to improving the performance of sports facilities by moving the cursor in the following scale, where 0 means the least effort, and 100 means the maximum effort. Briefly justify your choice here.
- Indicate how you would change the amount of budget assigned to sports facilities the next year, by selecting one of the following options (from +50% to -50%). Briefly justify your choice here.

Appendix B - Experimental scenarios, chapter 3

The text in italics displays the experimental manipulations. With respect to the control groups' vignette, one of the three sentences in italics in the square brackets is added to the other experimental groups' vignettes.

Experiment 1

Imagine you are the superintendent of a school district. You have to choose between two management software packages of the same price: either Sigma or Orion. Both software packages have been approved by the Ministry. However, there is no evidence of one being better than the other. [*The ministerial guidelines suggest the adoption of Orion.*] [*You heard that all of the school districts with the best national reputation will adopt Orion.*] [*The professional association of superintendents to which you belong suggests that you adopt Orion.*] Which software would you buy?

Experiment 2

Imagine you are the superintendent of a school district. You have to choose between two management software packages of the same price: either Sigma or Orion. Both software packages have been approved by the Ministry. Orion's performance is slightly worse than the other. [*However, ministerial guidelines suggest the adoption of Orion.*] [*However, you heard that all of the school districts with the best national reputation will adopt Orion.*] [*However, the professional association of superintendents to which you belong suggests that you adopt Orion.*] Which software would you buy?

Experiment 3a

Your organization must choose which management software to adopt between two options of the same price: either Sigma or Orion. Both software packages have been approved by the Ministry. Orion's performance is slightly worse than the other. [*However, ministerial guidelines suggest the adoption of Orion.*] [*However, you heard that all of the school districts with the best national reputation will adopt Orion.*] [*However, the professional association to which you belong suggests that you adopt Orion.*] Which software do you choose to adopt?

Experiment 3b

You must decide which training program to attend between the following two: either Delta or Gamma. The two programs provide the same number of credits, have the same schedule of classes, and require exactly the same effort. The performance

assessment issued by an entirely reliable independent agency indicated that Gamma is slightly worse than the other. *[However, the HR director of your organization suggests attending Gamma.] [However, your colleagues who all recognize as the best will attend Gamma.] [However, the people with whom you have shared your studies or your career path - people very similar to you - will attend Gamma.]* Which course will you attend?

Experiment 4a

Your organization must choose which management software to adopt between two options of the same price: either Sigma or Orion. Both software packages have been approved by the Ministry and are completely identical except for the features described as follows. In particular, the performance assessment issued by an entirely reliable independent agency indicated that Orion is slightly inferior to Sigma in terms of ease of use, speed, accuracy, and technical support. *[However, ministerial guidelines suggest the adoption of Orion.] [However, you heard that all of the school districts with the best national reputation will adopt Orion.] [However, the professional association to which you belong suggests that you adopt Orion.]* Which software do you choose to adopt?

Experiment 4b

You must decide which training program to attend between the following two: either Delta or Gamma. The two programs provide the same number of credits, have the same schedule of classes, and require exactly the same effort. The performance assessment issued by an entirely reliable independent agency indicated that Gamma is slightly worse than Delta for practical usefulness, quality of the instructors, quality of the content, and efficacy of the teaching methodologies. *[However, the HR director of your organization suggests attending Gamma.] [However, your colleagues who all recognize as the best will attend Gamma.] [However, the people with whom you have shared your studies or your career path - people very similar to you - will attend Gamma.]* Which course will you attend?

Experiment 5a

Your organization must choose which management software to adopt between two options of the same price: either Sigma or Orion. Both software packages have been approved by the Ministry and are completely identical except for the characteristics described in the table below. The table shows the performance assessment of the two software packages, on a scale from 0 (minimum) to 100 (maximum), issued by an entirely reliable independent agency.

	Sigma	Orion
Easy-to-use	74	72
Speed	89	87
Accuracy	93	91
Technical support	76	74

[However, ministerial guidelines suggest the adoption of Orion.] [However, you heard that all of the school districts with the best national reputation will adopt Orion.] [However, the professional association to which you belong suggests that you adopt Orion.] Which software do you choose to adopt?

Experiment 5b

You must decide which training program to attend between the following two: either Delta or Gamma. The two programs provide the same number of credits, have the same schedule of classes, and require exactly the same effort. The performance assessment of the two courses, on a scale from 0 (minimum) to 100 (maximum), issued by an entirely reliable independent agency is shown in the table below.

	Sigma	Orion
Practical usefulness	74	72
Quality of the instructors	89	87
Quality of the content	93	91
Efficacy of the teaching methodologies	76	74

[However, the HR director of your organization suggests attending Gamma.] [However, your colleagues who all recognize as the best will attend Gamma.] [However, the people with whom you have shared your studies or your career path - people very similar to you - will attend Gamma.] Which course will you attend?

Appendix C - Additional evidence supporting the coding, chapter 3

Additional evidence supporting the coding

Second-order themes	First order codes
COERCIVE ISOMORPHISM	
Attributes of decisions	<p><i>Based on reliable studies</i></p> <p>“Guidelines are based on reliable studies that are carried out in depth and that no single public agency could afford”</p> <p>“Guidelines are created starting from a wellknown casuistry”</p> <p><i>Drafted by experts</i></p> <p>“Civil servants should stick as much as possible to ministerial guidelines as these are drafted by experts in the sector”</p> <p>“Guidelines are drafted on the basis of the opinion of experts in the sector and civil servants should pay serious attention to them”</p> <p>“Guidelines should be followed by civil servants considering that guidelines are written by highly qualified public personnel”</p> <p>“Guidelines are written by experts with technical know-how and, as such, civil servants should make good use of them”</p> <p><i>Based on practical experience</i></p> <p>“Guidelines are drafted based on previous experience and therefore should be followed, or at least should accompany the civil servants when making decisions”</p> <p>“Guidelines perform the function of orienting towards services and standards that have been tested already and that can be therefore considered reliable”</p>
Decision support (process)	<p><i>Reference point for decisions</i></p> <p>“In general, guidelines represent a certainty”</p> <p>“Guidelines represent a useful tool of reference and consultation, thanks to which you can retrieve the information you need easily and behave accordingly”</p> <p>“Guidelines are important because they make life easier for whoever is called to make a decision”</p> <p><i>Enabler of professional conduct</i></p> <p>“Certainly, guidelines aim at preventing someone from going off the beaten track and this should be seen as a support, as an advice not to lose sight of the goal”</p> <p>“Guidelines are useful to define some constraints and it is always useful to set boundaries into what may otherwise be too vague as the behavioral expectations towards civil servants”</p> <p>“Guidelines enable the civil servant to be advised on her/his professional decisions”</p> <p>“Thanks to the guidelines, one can work without problems and even avoid being suited for nothing, as is sometimes the case in the public sector”</p>
Homogenization	<p><i>Guarantee of uniformity in the public sector</i></p> <p>“Ministerial guidelines tend to channel the action of the public sector in a common route of efficiency. This is why they should be considered and we should base our action upon them”</p> <p>“Guidelines should, in principle, harmonize the activities of the public sector”</p> <p>“I believe they are crucial to avoid that each agency operates according to its own way and to avoid useless duplications in looking for the same information”</p>
MIMETIC ISOMORPHISM (BEST PERFORMING COLLEAGUES)	
Attributes of decisions	<p><i>Based on serious and committed assessment by the “best colleagues”</i></p> <p>“The choices of the best colleagues are based on a cost benefit analysis”</p> <p>“The best colleagues decide based on criteria that keep in high consideration quality, and in this respect one should follow their choice”</p> <p>“Typically, best-performing colleagues display high levels of effort in choosing”</p> <p>“The best colleagues decide on the basis of evaluations resulting from the analysis of the content of, for example, training programs and of the professional level of the instructors and of the school”</p> <p>“The best colleagues choose based on sound evaluation criteria and perform well in their decisions—their advice should certainly be followed because it results from a serious and thoughtful evaluation”</p> <p><i>Based on practical applicability</i></p> <p>“The colleagues recognized by all as the best select on the basis of the practical utility of a specific option, such as a training program or of a software and yes, I believe one should follow their choice”</p> <p>“I think the best colleagues assess carefully a series of criteria when choosing, including the practical utility for the agency, and I believe it is useful to base the choice upon what the best say and do”</p>
Decision support (process)	<p><i>Support for those with scarce experience or skills</i></p> <p>“The best colleagues are trusted based on objective observation of their competency, professional approach, and commitment. Their advice can be especially useful for those who are not particularly experienced”</p> <p>“I believe the best can provide a useful decisional support to those who do not have the same experience and ability”</p>

Additional evidence supporting the coding, continued

Second-order themes	First order codes
Decision connected to the qualities of the “best colleagues”	<p><i>Decisions made by the “role models”</i></p> <p>“Their decisions are typically aimed at improving their professional profile and their agency, and this could serve as a role model for their colleagues”</p> <p>“The esteem and high considerations towards the best colleagues are great reasons to reflect upon what they choose and possibly do the same”</p> <p><i>Decisions made by the pioneers</i></p> <p>“I always think that the best colleagues understand earlier and better the validity of a specific training program or software, and I am therefore inclined to follow them”</p>
MIMETIC ISOMORPHISM (PUBLIC AGENCIES WITH THE HIGHEST REPUTATION) Attributes of decisions	<p><i>Based on the fit with the needs</i></p> <p>“The agencies with a good reputation spend time assessing the needs of employees and then select courses which should be followed or else each agency should go through the same assessment”</p> <p>“They choose on the basis of the needs of personnel and on the updates that are needed to keep up with progress and changes”</p> <p>“For the agencies that actually have a good reputation, I believe training is chosen based on a serious analysis of the real needs of the organization and of the personnel”</p> <p><i>Based on objective criteria</i></p> <p>“They choose based on the efficiency and the merit and you can always trust these principles”</p> <p>“They choose based on what can be considered universal criteria”</p> <p>“Certainly the primary criteria include costs, speed, reliability, and reputation”</p> <p>“Criteria of efficiency and effectiveness”</p> <p>“Based on criteria of efficiency and effectiveness, which are objective”</p> <p><i>Based on a careful assessment</i></p> <p>“They choose based on cost-benefit analysis”</p> <p>“They normally analyze precisely the relation between costs and benefits”</p> <p>“The criteria employed correspond to a careful assessment of the provider”</p> <p>“The agencies with the best reputation are overall those that are virtuous in spending public money, therefore they will choose maximizing quality vs cost”</p>
Decision support (process)	<p><i>Specifying criteria for selection</i></p> <p>“Criteria for selection are qualitative and quantitative, these are like standards for the rest of us”</p> <p>“If you look at the best agencies and follow them you can’t be wrong. At the end of the day, they have built their reputation with commitment and sacrifice and they select on the basis of their know-how that is then shared”</p>
Decision connected to the qualities of the “best public agencies”	<p><i>Decisions made by the pioneering agencies</i></p> <p>“The best public agencies are those that implement first innovation and reform processes and that display efficiency. I believe their decisions should be trusted because they are based on experience and on serious motivations”</p> <p>“They select based on their needs but also on the basis of the scientific, technological, and organizational progresses”</p> <p><i>Decisions made by the trusted agencies</i></p> <p>“Certain agencies have a great reputation and deliver great services—they choose on the basis of what keeps their employees always updated and therefore should be followed”</p>
NORMATIVE ISOMORPHISM Attributes of decisions	<p><i>Decisions aimed at developing competences</i></p> <p>“Professional associations inform, train, guarantee, and offer guidelines to their members, and they provide suggestions that are in line with this purpose”</p> <p>“I believe their role is to offer services useful for the development of the profession, such as training, updates, legal counseling, and when they recommend something all these criteria should have been included in their decision”</p> <p><i>Decisions aimed at protecting members’ rights</i></p> <p>“The role of a professional association is to protect its members and their rights, suggesting solutions that are in line with this purpose”</p> <p>“The main role of professional associations is to defend, yes I mean defend, the interests of their members. I don’t know if they always choose or recommend the best in absolute terms, but certainly they choose based on the protection of their members”</p> <p>“If you follow the guidelines of professional associations, you know they will guarantee you from a legal point of view. There can’t be disputes, and even if there are disputes, the choice of the association will give protection from any legal procedure”</p>
Decision support	<p><i>Providing support and safe advice</i></p> <p>“The primary role of professional associations is to support and to advise. I follow them because their decisions and recommendations are safe”</p> <p>“This is the place where professional advice is given, but also training and what allows to improve civil service from a professional point of view”</p>
Homogenization	<p><i>Guarantee uniformity of conduct among professionals</i></p> <p>“Professional associations unite people doing the same job”</p> <p>“In my view, the role of professional associations is to homogenize and to blend the associates around a certain subject or project. Therefore, I expect associates to conform to what in a sense they have decided together”</p> <p>“The role of professional associations, their actions and their choices, is to guarantee seriousness and professionalism to a category of professionals”</p> <p>“Following the recommendations of professional associations helps maintain a certain internal conduct within the category and this will also be recognized outside”</p> <p>“Their main role is to offer a specialized reference point for professionals and this includes guarantees but also guidelines”</p>

Appendix D - Additional evidence supporting the coding: ex-post explanations, chapter 3

Additional evidence supporting the coding: ex-post explanations

Second-order themes	First order codes
COERCIVE ISOMORPHISM	
Reasons for following ministerial guidelines that recommend an inferior solution	<p><i>Trust towards the decision makers</i></p> <p>"I trust what the ministry recommends, no matter what"</p> <p>"I guess if you trust an agency and its experts, you don't question much"</p> <p><i>Alternative dimensions of performance</i></p> <p>"Guidelines may recommend a training program or a software with inferior performance in case of cost cuts if the worse is cheaper"</p> <p>"I would like to understand why the ministry is recommending a worse solution but, in general, if something is recommended by the ministry I tend to follow it. It may be, for example, that the solution is tailored to the workload"</p> <p>"Perhaps they recommend an inferior training or software because, while being slightly inferior, it better meets the needs of the agency"</p> <p>"One should consider that the best product is not always the winning one—history is packed with examples of products qualitatively inferior that then have become standard and therefore have won"</p> <p><i>Homogenization</i></p> <p>"It may be that there are constraints of homogeneity with the existing solutions—then I would follow the ministerial guidelines"</p> <p>"I think I would follow the advice of a worse-performing solution, also because otherwise I may run the risk of misalignment with the great majority of the colleagues. Therefore, the supposed 'best performance' of a certain solution would end up being useless or even counterproductive"</p> <p>"I don't have a clue about the reason why the ministerial guidelines may recommend a worse product. I would follow them in any case to feel safer and more guaranteed"</p>
MIMETIC ISOMORPHISM (BEST PERFORMING COLLEAGUES)	
Reasons for following ministerial guidelines that recommend an inferior solution	<p><i>Trust towards the best colleagues and their experience</i></p> <p>"Yes, I would trust their recommendation if I trust them"</p> <p>"I would follow their advice to attend the training program or buy the software with inferior performance, because I start from the assumption that the best colleagues are more knowledgeable than me. I trust them"</p> <p><i>Alternative dimensions of performance</i></p> <p>"Sure, I would follow their advice because the fact that the performance is inferior doesn't mean these solutions are not useful"</p> <p>"These colleagues may have tried the program or software they recommend and therefore may have realized it works well, even better than the one which is in principle the best"</p> <p>"I would follow their advice because I expect it to be derived from their direct experience, which may have showed some elements that escape the general criteria of performance"</p> <p><i>Homogenization</i></p> <p>"In order to have an homogeneous training or IT program, I would follow their advice"</p>
MIMETIC ISOMORPHISM (PUBLIC AGENCIES WITH THE HIGHEST REPUTATION)	
Reasons for following ministerial guidelines that recommend an inferior solution	<p><i>Trust towards the best agencies</i></p> <p>"The agencies with the best reputation are normally those you can trust. If they choose something, there must be a reason"</p> <p><i>Alternative dimensions of performance</i></p> <p>"I would follow their advice. Probably, they have chosen the worse-performing one to prevent problems in the future, if they know the provider is not doing great and investing in innovation. Today it may be the best solution right now, but not in a year or so."</p> <p><i>Homogenization</i></p> <p>"I would follow their advice because the solution they opt for may be likely to be the most common among all agencies"</p>
NORMATIVE ISOMORPHISM	
Reasons for following ministerial guidelines that recommend an inferior solution	<p><i>Trust towards the professional associations</i></p> <p>"I would follow the professional association. I trust them, they protect my interests, and I will do what they tell me to do"</p> <p>"Probably they have their good reasons and I would stick to their advice"</p> <p><i>Alternative dimensions of performance</i></p> <p>"There may be many reasons to choose an inferior product, not the least the fact that it is more adaptable to the needs of a specific professional category"</p> <p>"Yes, a professional association has all the skills to choose and advice, it is not up to me to decide whether it is inferior or not. The most important thing is the outcome and a better knowledge of the problems may allow them to choose the solution that apparently performs worse but that achieves better results"</p>