Article

Decision Biases and Heuristics Among Emergency Managers: Just Like the Public They Manage For?

American Review of Public Administration
2019, Vol. 49(3) 292–308

DOI: 10.1177/0275074018799490
journals.sagepub.com/home/arp
SAGE

Patrick S. Roberts[1] **and Kris Wernstedt**[1,2]

Abstract

We present evidence that emergency managers exhibit some of the same decision biases, sensitivity to framing, and heuristics found in studies of the general public, even when making decisions in their area of expertise. Our national survey of county-level emergency managers finds that managers appear more risk averse when the outcomes of actions are framed as gains than when equivalent outcomes are framed as losses, a finding that is consistent with prospect theory. We also find evidence that the perceived actions of emergency managers in neighboring jurisdictions affect the choices a manager makes. In addition, our managers show evidence of attribution bias, outcome bias, and difficulties processing numerical information, particularly probabilities compared to frequencies. Each of these departures from perfect rationality points to potential shortfalls in public managers' decision making. We suggest opportunities to improve decision making through reframing problems, providing training in structured decision-making processes, and employing different choice architectures to nudge behavior in a beneficial direction.

Keywords

behavioral public administration, emergency management, decision making, heuristics, prospect theory

Behavioral economics and cognitive psychology have produced insights about the limits of human rationality in decision making. Recognition of the prevalence of decision biases and heuristics among the public led to the "nudge" concept—the idea that bureaucrats can make minor changes to choice architecture to improve people's behavior (Thaler & Sunstein, 2003, 2008). One example is a government agency that tells doctors that they are unusually high prescribers of controlled substances to encourage them prescribe fewer such substances because their decision biases lead them to want to go along with the crowd. The biases and heuristics employed by the public—the ones that the nudge aims to correct—are well studied. The bureaucrats who employ the nudge are less well-studied, and their own biases and mental shortcuts are not well documented. Anecdotal evidence and conversations with colleagues at professional conferences imply a common recognition that such decision biases and heuristics likely exist among bureaucrats but, as we discuss below, few studies have taken the important step of actually documenting them. Therefore, we ask whether decision biases and heuristics exist among one group of public managers—emergency managers. We also ask whether knowledge about decision biases and heuristics can improve emergency management and public management more generally.

The study of biases and heuristics has proceeded along interlocking paths in psychology and economics. Both disciplines attempt to document features in individual and group cognition that lead people to make choices that, in the language of neoclassical economics, are departures from the utility-maximizing rational model, or, in psychology, that produce subpar outcomes. Decision biases are not normatively bad, as in the term "racial bias." Rather, they are departures from the model of how economists and psychologists traditionally have assumed that people made decisions.

In the study of public administration, identifying managerial biases and decision shortcuts holds out promise for improving understanding of the decision-making process. A burgeoning research program in behavioral public administration uses theories from psychology to test the microfoundations of administrative behavior at the individual level (Grimmelikhuijsen, Jilke, Olsen, & Tummers, 2017; Moynihan, 2018). Attention to individual-level biases among public administrators is a largely unexplored territory, however, with some exceptions, including work

[1]Virginia Tech, Alexandria, USA
[2]Ardhi University, Dar es Salaam, Tanzania

Corresponding Author:
Patrick S. Roberts, Associate Professor, Center for Public Administration and Policy, School of Public and International Affairs, Virginia Tech, 1021 Prince Street, Alexandria, VA 22314, USA.
Email: robertsp@vt.edu

on heuristics employed by fire managers (Arvai, Gregory, Ohlson, Blackwell, & Gray, 2006; Zaksek & Arvai, 2004), and more general work on risk perceptions of transportation officials (Feeney & Smith, 2008), local solid waste planners (Snary, 2004), and city managers (Rahm & Reddick, 2011). Understanding how bureaucrats make decisions extends a research tradition dating to at least Simon (1947), and it offers the promise of using a better understanding of decision making to develop better decision processes.

We attempt to identify whether public managers exhibit the same decision biases and heuristics found in other studies of the nonexpert general public through a national survey of county-level emergency managers. We study these particular bureaucrats because we want to examine decision making by individuals who are highly experienced with making decisions under conditions of risk and uncertainty. This allows us to present our study group with decision scenarios that explicitly match their field of expertise, which is more specific than the stylized behavioral economics or psychology problems often presented to classrooms of college students or Mechanical Turk samples.

County-level emergency managers provide fascinating empirical ground for observing biases and heuristics because they operate as an important fulcrum in the system of disaster management. In normal times, they engage in planning and mitigation efforts with state and local government agencies, nonprofits, and citizen groups, to reduce the damage caused by disasters. They are supposed to identify the range of hazards in a community including fires, floods, and industrial accidents, and then develop programs to address those hazards, in cooperation with other government actors. In the case of floods, emergency managers develop plans to mitigate the effects of heavy rainfall and streamflow before they happen. Biases and heuristics may shape these decisions. For example, managers might follow the lead of a neighboring county without taking the time to consider all relevant facts, a phenomenon we call the neighbor effect. They also might allocate more resources when an action to prevent damage is framed as potentially preventing a loss than when it is framed as potentially securing a gain. Both of these options would be consistent with prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992), a well-known concept from psychology. There's nothing wrong with taking extra precautions but emergency managers are charged with using collective resources prudently and fairly.

Why Document Biases and Heuristics Among Bureaucrats?

We expect that public managers will exhibit some of the same biases and heuristics as the rest of the population (Kahneman, Knetsch, & Thaler, 1991). Is this simply proving the obvious? Is it "duh" research? We think not. Documenting the existence of these biases and heuristics is important for three reasons.

Evidence Matters

Even if it is commonly assumed that public sector managers exhibit the same biases and heuristics as the general public, we should offer evidence to support that assumption. Some scholars argue that the biases and heuristics identified by behavioral economics are not as strong as is sometimes presented (Wright, 2007), or result from a weak research design (Plott & Zeiler, 2007). Other scholars assume that politicians exhibit the same biases and heuristics as the general public, without providing empirical support (McDermott, Fowler, & Smirnov, 2008; Mercer, 2005). Still others assert that "bureaucrats are far from free of cognitive biases of their own"—an assertion that may be correct but that lacks systematic empirical support (Somin, 2017). Sunstein (2016, p. 76) offers a litany of decision biases that afflict bureaucrats, and writes that the field of behavioral public choice is "rapidly growing, and it is likely to prove highly productive." He offers no empirical support for which biases bureaucrats are likely to have, however. The work he cites by Schnellenbach and Schubert (2014, p. 28) to support his claim finds that, "the choices of bureaucrats and lobbyists are to a large extent still uncharted territory."

Our article contributes to this debate by providing evidence of biases and heuristics among emergency managers in their domain of expertise. Anticipating objections to this seemingly narrow target group, our focus on county-level emergency managers allows us to provide a highly relevant and well-understood decision context to study participants whose important role in reducing disaster risk has long been a focus in public administration (Comfort, Waugh, & Cigler, 2012). In contrast, including a wider range of managers would require us to use a very general treatment that has less relevance for the specific decisions that different types of managers make.

Trust Us, We're the Experts?

Studies of expertise in economics, business, and chess suggest that experts make decisions in more strategic or rational ways than do nonexperts, and, therefore, the effects of framing and biases may be attenuated (Druckman & McDermott, 2008; Hafner-Burton, Hughes, & Victor, 2013). For instance, research in the weather forecasting domain suggests that experts who regularly work with uncertain information have better numeracy skills and a better ability to interpret probabilities than the general public (Murphy & Winkler, 1977; Stewart, Roebber, & Bosart, 1997). The assumption is that the experts know better because they can decide better. Our analysis of biases and heuristics among experts adds nuance to assumptions and claims that experts know better—they may know better, but

they also fall short of rational decision-making ideals in particular ways that we identify empirically.

Our emergency managers qualify as experts because they have training and experience in preparing for floods, and are recognized as an expert in their counties. While the nature of expertise is contested, most definitions include domain-specific competencies and a long period of repeated performance (Ericsson et al., 2018, pp. 3-4). Mieg (2006, p. 743) recommends that an investigation begin with the study of an expert in his or her functional role. Bureaucratic expertise also connotes a recognized claim to authority, which is distinct from but often related to skill and performance (Bendor, Taylor, & Van Gaalen, 1985; Rose, 1993).

Public Managers Could Be Nudged, Too

The literature on nudges designed to correct decision biases neglects accounting for the fact that bureaucrats who provide the nudge may themselves be subject to decision biases (Thaler & Sunstein, 2003, 2008). In 2015, President Obama's executive order 13707 required federal agencies to integrate behavioral insights (another word for "nudges") into their work. The intent was for bureaucrats to alter the choice architecture of citizens to achieve a widely recognized good, but the program begs the question of whether bureaucrats themselves could be subject to a nudge in cases where there is an objectively better decision.

For all of these reasons, we collect evidence to determine whether one type of public administrator is subject to decision biases by presenting a series of hypothetical decision scenarios to a sample of emergency managers and asking the managers to choose their preferred options. These hypothetical situations obviously cannot capture all of the stresses, accountability, and other real-world nuances of an actual decision, or even those features that a naturalistic decision-making study might examine (Klein, 2008). However, the other scenarios usefully allow us to work with a real-world set of decision makers about problems in their domain of expertise to test limits to rationality across an array of heuristics that we describe next.

Hypotheses

Prospect Theory and the Emergency Management Environment

We start our tour of public administrators' heuristics by testing some of the implications of prospect theory. Most public administration scholars have been exposed to this concept since it appeared in the economics literature nearly 40 years ago and permeated most other decision sciences since, albeit it remains discussed almost exclusively in the context of decision making by business executives and the general public rather than government managers (Kahneman & Tversky, 1979; Mercer, 2005; Tversky & Kahneman, 1981) with recent exceptions (i.e., Baekgaard, 2017). In brief, prospect theory departs from the standard utility-maximizing model of rational choice by beginning with the assumption that the different psychological impact of gains and losses shapes decision calculations. In particular, people appear more risk averse in the domain of gains, and more risk seeking in the domain of losses. As a simple example, prospect theory predicts that the loss in satisfaction from losing 10 units of some good exceeds the gain in satisfaction from gaining 10 units (Abdellaoui, Bleichrodt, & Paraschiv, 2007; Kahneman & Tversky, 1979). An implication of this is that people will more readily gamble in situations to avoid losses than in situation that involve potential gains. They also make their risk calculation based on the marginal change they would be experiencing in satisfaction and not their absolute level of satisfaction (the so-called "reference dependence" effect). In addition, Kahneman and Tversky (1979) find that individuals underweight probable outcomes as opposed to certain ones. The upshot of these features is that individuals in the general public, and maybe even expert managers accustomed to having to deal with uncertainty, have trouble interpreting uncertainty and probabilistic information. This also means that the framing of an uncertain situation can influence the risk calculation and decision that a manager makes.

Among the tasks county-level emergency managers face are deciding upon plans to prepare for floods, recommending how to allocate resources to protect different geographic areas, implementing a plan over a fast or slow timeline, and spending county resources. Cooper and Kovacic (2012) propose that managers face a trade-off between maximizing social welfare and satisfying their career imperatives by pleasing politicians, and that they, therefore, feel political pressure to obey fiscal constraints. Emergency managers themselves report concern about their reputation in interviews, and media coverage after costly disasters sometimes places the blame on them (Payne & Morris , 2016). Moreover, general management and emergency management-specific theory hold that managers are conscious of their "reputational capital," because it represents an important currency in getting things done across networks with little formal control (Koliba, Mills, & Zia, 2011; Kreps & Wilson, 1982; Morris, Morris, & Jones, 2007, p. 95; Wong & Boh, 2010). Thus, the emergency manager's decision environment includes calculations about minimizing loss of life and property in their jurisdiction as well as about protecting their reputation by calling for extra protection against a large disaster that eventually occurs.

We apply prospect theory to emergency management by proposing that emergency managers will approach equivalent forecasts framed in the domain of gains differently than those framed in the domain of losses.

1. Prospect Theory Hypothesis: Emergency managers are more likely to take a risk in the domain of losses than when an action is presented as a gain.

Regret After Errors of Omission and Commission

Managers sometimes need to consider committing to an action that could help to minimize damages from a possible upcoming disaster, but that would incur some avoidable costs (financial and/or reputational) if the disaster fails to occur. The alternative would be to omit taking action to avoid risking an unnecessary mistake, at the cost of not adequately preparing for the disaster if it does occur. If the state of nature turns out to run opposite to the action taken, the manager needs to live with the consequences. We call these errors of commission and errors of omission, respectively. The terms "false positive" (a false prediction that an event will happen when the event does not actually occur) and "false negative" (a false prediction that an event will not happen when the event actually does occur) also apply. A nuclear plant operator looks at an unreliable warning signal, for instance, and needs to consider whether to ignore the warning or to respond to it. If she ignores it, perhaps a chain reaction will ensue, causing a loss of control and an overcooling accident (Nuclear Regulatory Commission, 1986). If she responds to it, perhaps the plant will needlessly shut down. More routinely, in health care, doctors and nurses may miss a warning sign of a disease, or may overreact to a false signal. Scholars have examined similar phenomena in terms of the well-known Type I and Type II statistical error classifications (Gill & Meier, 2000; Heimann, 1993; Landau & Stout, 1979).

With few exceptions, however, the literature has not dealt systematically and directly with how managers approach these different types of errors or signals in practice. Eiser et al. (2012) borrowed from signal detection theory (Swets, 1973) to discuss the generic problem in disaster management of discriminating whether information provides a signal to take a feasible action or just "noise" that lacks practical implications for the range of actions that could be taken. They note that the costs of misses or errors of omission can be catastrophic in such settings, suggesting a precautionary approach—and thereby risking an error of commission/false positive/Type I error—and also point out that the costs of false alarms can be large as well. These include the direct costs of an action, indirect opportunity costs, or the costs associated with heightening a cry wolf syndrome.

The emergency manager's relationship to these errors in decision making is particularly interesting because she is expected to be an expert in her domain. Citizens, government agencies, and elected city and county leaders rely on the emergency manager's judgment about how to prepare for disasters. However, if emergency managers have systematic biases, then the entire system of preparing for disasters may be influenced in the direction of the biases. With the potential for severe consequences in some relatively rare, extreme events—such as the loss of life from being swept away and drowning in a flood—it may seem imperative to avoid the sins of omission, false negatives, and Type II errors. While this situation does not completely align with the context of prospect theory—where the difference in risk preference in gain and loss contexts is defined in a context of equivalent expected values in decision outcomes—prospect theory does suggest that individuals tend to overweight the low-probability tail of extreme losses in their decision making.

2. Regret Hypothesis: Emergency managers have less regret about being wrong when the negative outcome is the result of an omission or failure to act than when the negative outcome is the result of an action they took (Baron & Ritov, 1994).

Attribution Bias

Research on decision making under uncertainty has demonstrated the presence of a so-called attribution bias, where one assesses the decision choices of others differently than one's own choices (Heider, 1958; Kelley, 1967; Ross, 1977). This reflects a tendency to amplify the role of idiosyncratic personality in explaining the behavior of others, and diminish the situational context of the behavior, and to reverse this for one's own behavior. This can lead to a number of observed biases in group decision making common in public management, such as false consensus (Krueger & Clement, 1994). It also can exacerbate self-serving biases for individuals to remain immune to constructive negative feedback.

3. Attribution Bias Hypothesis: Emergency managers assess the decisions of others differently than their own decisions (Heider, 1958; Kelley, 1967; Ross, 1977).

Numeracy

Scholars have found that many people better at processing natural frequencies, expressed in statements such as 9 in 10, than at processing conditional probabilities, which are often represented as simple percentages or proportions such as 90% or 0.9 (Gigerenzer & Hoffrage, 1995, p. 697). Cosmides and Tooby (1996) theorize that natural frequencies offer greater computational simplicity and are more familiar to the human mind as it evolved, with probabilities having arrived on the scene later in the evolutionary timescale. However, the processes undergirding numerical processing remain opaque (Gigerenzer, 1994; Reyna & Brainerd, 2008; Slovic, Monahan, & MacGregor, 2000). Research continues on the impact of presentations of frequencies versus probabilities, but most of this work surveys the general public or medical professionals, not public managers. While the thrust of the literature is that many people more quickly and accurately understand frequencies, data are most often presented to public managers in terms of probabilities.

Prior research suggests a shortfall in numeracy skills can exacerbate problems with framing risk and uncertainty (Peters, Västfjäll, Gärling, & Slovic, 2006). For instance,

those with more limited numeracy skills are more likely to rate beef that is 75% lean more favorably than beef that is 25% fat, and view the risk of a bad outcome that occurs in 10 out of 100 cases with more concern than the risk of a bad outcome that occurs in 10% of cases. In addition, the intensity of feelings that individuals may associate with particularly dreadful or pleasurable events influences the ability to make optimal decisions with probabilistic information.

We do not have enough prior literature in the management context to predict exactly how managers will react to forecasts presented in different forms. However, the prior literature does suggest that our emergency managers may show different risk preferences to mathematically equivalent forecast information depending on the form of that information.

4. Numeracy Hypothesis: Emergency managers exhibit different risk preferences depending on the format in which otherwise identical information is presented.

Neighbor Effect

The terms groupthink, bandwagon effect, and herd effect refer to situations in which people are more likely to follow an apparent consensus. Groupthink occurs when people refrain from expressing opinions because of the desire to conform to a perceived consensus (Janis, 1982). When people follow cues about popularity and "jump on the bandwagon" of a status symbol good or social movement, they exhibit evidence of the bandwagon effect (Kastanakis & Balabanis, 2012).

Most social psychologists prefer to analyze what they call the herd effect or herd behavior—situations in which people act based on the decisions of others (Rook, 2006). Scholars try to tease out whether people follow the lead of those around them as a response to new information or whether they copy their behavior as a psychological tool of imitation, designed to alleviate anxiety caused by uncertainty (Chen, 2008; Christie & Huang, 1995; Shiller, 1995; Thaler & Sunstein, 2008, pp. 53-73). Some studies find that the herd effect is a rational means of spreading information that helps systems to adapt to new data—which raises the question of whether the effect is a decision bias in the sense of a feature that does not maximize rational utility (Devenow & Welch, 1996; Zhao et al., 2011).

Other studies root the herd effect in a psychological response to imitate others as a way to cope with the fear of making incorrect decisions under conditions of uncertainty (Christie & Huang, 1995). Thaler and Sunstein (2008, p. 56) cite a famous study by Solomon E. Asch (1956) in which respondents answer questions in a way that is clearly false, just because they want to go along with the crowd. The questions include trivial items such as saying that two lines are of equal length when they clearly are not. In a groupthink-oriented study, Thaler and Sunstein (2008, pp. 54-60) provide examples of people who go along with the group because of peer pressure, and then show much less conformity when decisions are anonymous. This difference suggests a social dimension to the herd behavior rather than a purely informational one.

In the emergency management context, managers may take different actions depending on what peers in neighboring jurisdictions are doing, even when other elements of the decision problem do not change. An emergency manager might learn that a counterpart in a neighboring county ordered an evacuation or implemented an emergency management plan, and then decide to do the same. This represents a phenomenon similar to the herd effect, although in our situation it is more of a pseudo-herd-effect or, more usefully, a "neighbor effect." We cannot measure the effect of imitation as separate from other reasons a manager might adopt the same decision as a counterpart. Instead, we investigate a practical situation to determine whether an emergency manager is more likely to act based on the decisions of managers in neighboring counties.

5. Neighbor Effect Hypothesis: The actions that emergency managers take are influenced by the actions that peers in neighboring jurisdictions take, even when other elements of the decision problem do not change (Chen, 2008; Shiller, 1995; Thaler & Sunstein, 2008, pp. 53-73).

Outcome Bias

Outcome bias occurs when people rate the decision quality as better, or the decision maker as more competent, when the outcome was favorable than when it was unfavorable (Baron & Hershey, 1988; Hawkins & Hastie, 1990). Emergency managers must make decisions in the face of risk and uncertainty (Moynihan, 2008). They may have to choose a plan to prepare for a flood although they do not know the timing, extent, or severity of the eventual flood.

6. Outcome Bias Hypothesis: We expect that emergency managers will rate a decision as better when the outcome was favorable than when it was unfavorable.

Method and Sample Characteristics

To test for the presence of decision biases and heuristics among emergency managers, we first had to find our sample. The most common subjects for studies of decision making under uncertainty are college students, but researchers have challenged the idea that students are acceptable proxies for the general public or for more expert decision makers (Fatas, Neugebauer, & Tamborero, 2007; Kinder & Palfrey, 1993; Plott, 1982).

Rather than college students or the general public, we chose to study the more challenging subject pool of emergency management professionals. For this, we compiled a list from public online sources of readily identifiable county-level

Table 1. Questionnaire Design.

Question	Question topics
1–8	Questions on background of respondent (e.g., age, gender, education, professional setting, years of professional work, experience with flooding)
9-27	Eight experiments or groups of questions (ranging from one to five questions per experiment). Each experiment has four hypothetical scenarios, with each respondent randomly sorted into one scenario. In this article, we report on results from Experiment 1 (gain and loss preferences), Experiment 2 (errors of omission and commission), Experiment 4 (decision process evaluation), Experiment 5 (ambiguity in probability expression), and Experiment 6 (numeracy).
28-30	Questions regarding the risk preferences of respondent (not reported in article).
31-38	Questions related to the relative importance of aspects of respondent's work, such as the protection of various public resources and job concerns (not reported in article).
39-42	Questions regarding the relationships of respondent with colleagues and the public (not reported in article).
43-46	Question on the respondent's comfort with decision making under uncertainty and the level of training provided.

Table 2. Sample Characteristics.

	% of sample (n = 316)[a]	% of Weaver et al. (2014) sample	IAEM (2005) sample
Male	80.1	80.9	76.1
45 years or older	84.8	71.0[b]	75.6[c]
10 or more years of work experience	68.0	71.2	—
Completed graduate or professional degree	17.1	23.4	38.1
Work mostly in rural areas	47.8	46.4	—
Flood has occurred in area within last 10 years	76.4	67.1	—

Note. IAEM = International Association of Emergency Managers.
[a]n = 276 respondents for the variable related to rural-nonrural location.
[b]46 and older in Weaver *et al.* (2014).
[c]40 years and older in IAEM sample.

emergency managers in each state in the United States, which we then stratified so as to include overrepresentation of rural managers in our sampling frame to compensate for expected lower response rates.[1] Our rural areas represent noncore counties, as defined by the Office of Management and Budget.[2] Rural managers may rely less on colleagues within their offices and counties and more on external networks for interpreting forecast information than managers in urban counties with larger staff (Roberts & Wernstedt, 2016). Importantly, our frame does not represent an equal probability sample as we would expect some variation between emergency managers' characteristics and the availability of readily available online contact information for them.

Prior to administering our online questionnaire, we pretested versions with six students and four local emergency managers. After revising the questionnaire in response to their feedback, principally by shortening its length, we implemented it as a pilot to 60 emergency management professionals not already part of our sampling frame to test its administration. This motivated some minor changes in our recruitment protocol.

The final version of our survey questionnaire contains 46 closed-ended questions (see Table 1), some of which relate to factual background information—the age, gender, and experience of respondents, for example—and others to hypothetical emergency management scenarios with flood and climate forecast uncertainty that we use to elicit information about the decision process in light of risks. These scenarios or survey experiments, parts of five of which we discuss here, differ in the details of their decision context. Recent studies provide empirical support for the external validity of such vignette-based experiments, showing that the decisions they elect do reflect real-world decisions (Evans et al., 2015; Hainmueller, Hangartner, & Yamamoto, 2015; Peabody et al., 2004). The possible question responses themselves were either binary (option 1 or option 2) or Likert-type or other ordinally scaled answers indicating relative agreement or preference. For this reason, we use tests of proportions and nonparametric tests to investigate each of our six hypotheses as described below.

Our online survey platform (Qualtrics) allows us to randomly sort respondents into one of four versions in each experiment; that is, we break respondents into four groups for Experiment 1, four groups for Experiment 2, and so forth. The composition of each group varies across experiments[3] and each experiment has multiple questions (for that reason, they more precisely constitute experimental "sets"). Our final set of 19 questions focuses on general preferences and concerns.

Our survey administration entailed a standard two-stage approach. We first circulated over 1,600 email invitations to individuals in our sampling frame. Nonworking email

addresses returned from our first round of email invitations left us with 1,349 county-level emergency managers in the United States who plausibly received our invitation to complete the online questionnaire. We then followed up this initial invitation with two rounds of email reminders. The two steps yielded 316 respondents, of which 231 completed all questions. Our 18.7% response rate (American Association for Public Opinion Research, 2016, Survey Outcome Rate Calculator 4.0, Response Rate #4 definition) exceeds those (9%-13%) reported in the annual surveys conducted by the Federal Emergency Management Agency from 2011 to 2012 (see reports via www.fema.gov/local-official-survey-findings-flood-risk, accessed 7/19/2018). However, it falls below that of Weaver et al. (2014), who report a 30.3% response rate in their survey of more than 1,000 emergency managers. Our relatively lower response rate may result from including more challenging decision-scenarios in addition to the simpler background demographic and job-related questions found in the Weaver et al. study.

How well do our responses match the population as a whole? Unfortunately, no single list of the population of county-level emergency managers exists to answer this question. Each county structures its emergency management office independently from state or federal authorities, and some counties may have a half time emergency manager, and a small number may have none, while other counties may have two, or in urban areas an office of dozens. Sometimes the emergency management office is a standalone office, and sometimes it is combined with other services such as police, fire, or 911 emergency offices. Lacking population-level metrics, the best test we have is to compare our sample characteristics with other survey samples that attempt to capture the characteristics of emergency managers.

With this in mind, the characteristics of our respondents largely follow those surveyed in the Weaver et al. (2014) study, which is the most comprehensive survey of emergency managers available (Table 2). In both samples, a majority of respondents are 45 years and older, with more than 10 years of work experience, and have experienced a flood within the last 10 years. In addition, less than a third of respondents in each sample have received a graduate or professional degree, and in each, nearly half work in rural areas. Our sample demographics also resemble samples of emergency managers surveyed by Jensen and Youngs (2015), the International Emergency Management Association (International Association of Emergency Managers, 2005) and Peerbolte and Collins (2013), who similarly report a majority of older, educated, experienced males in the emergency management profession. In terms of geographic diversity, our respondents come from 29 states and all four U.S. census regions (Northeast, Midwest, South, and West). Emergency managers from the Midwest appear over-represented in terms of the Midwest's share of the more than 3,000 counties in the United States (53% of respondents vs. 34% of counties) and those from the South underrepresented (14% vs. 45%). Weaver et al. respondents show a similar underrepresentation of emergency managers from the South, although less sharp (33% vs. 45%).

Results

Prospect Theory Hypothesis

We divided our sample into four groups and exposed each group to a different scenario in which respondents assume they were an emergency manager in a fictional county and received a flood forecast. They then had to choose one disaster plan. One group of 83 randomly assigned emergency managers faced a choice between a plan that would result in the destruction of 75 out of 100 houses at risk and one that would have a 75% chance of resulting in the destruction of all 100 houses and a 25% chance of resulting in the destruction of 0 houses. As the right-most column in Figure 1 shows, 89% of the group's emergency managers chose the risk-preferring alternative that had a chance of resulting in 0 houses destroyed, and also a chance of destroying all 100 houses at risk. With another randomly selected group of 84 emergency managers, we used a gain frame to present one alternative of saving 25 out of the 100 houses with certainty (risk averse) versus the alternative to possibly save all 100 houses at the risk of saving none. Only 42% of this group's emergency managers chose the risk-preferring alternative of possibly saving all 100 homes (the second column from the left in Figure 1). Using a two-sample test of proportions, the responses show a statistically significant difference between the two frames; that is, the majority show risk aversion in the "gain" frame and a majority show risk preferring in the "loss" frame.[4] Our emergency managers, although more experienced with uncertain flood forecasts than the general public, appear sensitive to the same framing effects found in studies of prospect theory among student or general public populations (Kahneman & Tversky, 1984; Levin, Schneider, & Gaeth, 1998). If expert emergency managers are sensitive to framing effects, then the framing of forecast information becomes critical. Loss frames may encourage emergency managers to take more risks, and gain frames may encourage them to act more cautiously.

Regret Hypothesis

Our examination of how managers approach errors of omission versus commission relied on a flood forecast experiment in which our respondents answered a series of questions related to the impact of their decisions in response to that forecast. We told one group of 66 respondents to assume that they had decided to take action to reduce the risks associated with the potential flood, but the flood did not occur (error of commission). We told another group of 67 respondents that they had decided not to take

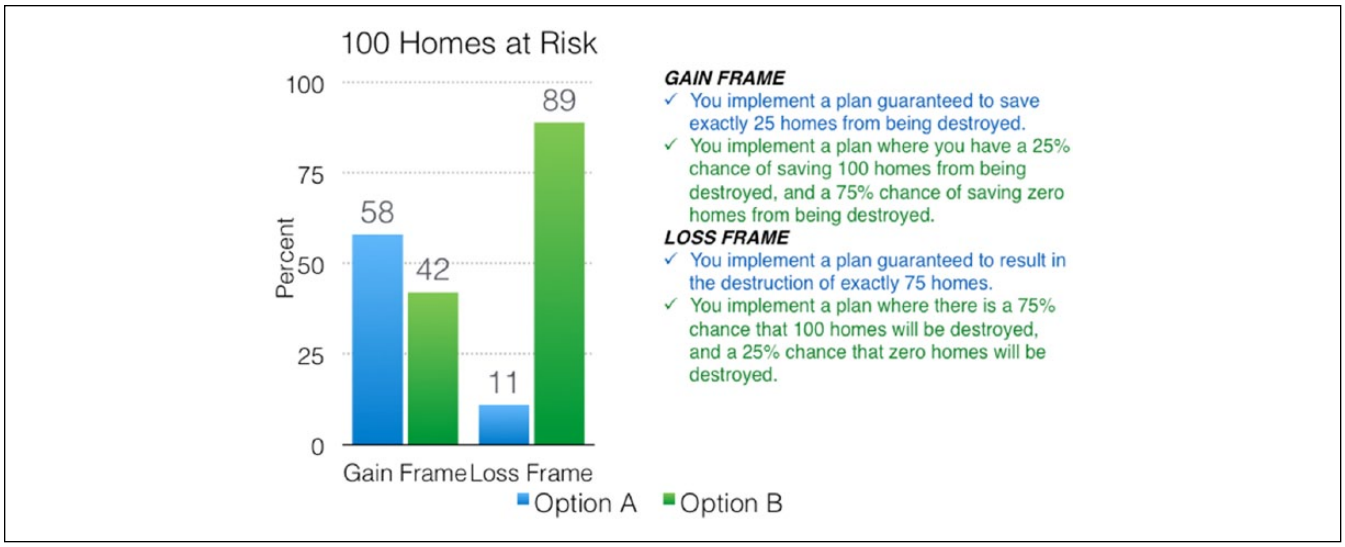

Figure 1. Prospect theory.

action to reduce the risks associated with the potential flood, but the flood did occur (error of omission). We then asked each group how concerned they were about their job security because of their error, on a scale of 1 ("extremely concerned about the security of my job") to 7 ("extremely confident about the security of my job"). This job security question gets at the cost of developing a poor reputation from making one of these errors. Emergency managers are vulnerable to being fired, or to having their departments reorganized and effectively being demoted after a destructive flood (Giwargis, 2017). They are also vulnerable to being fired by a vote of county supervisors, for reasons that include having spent money for flood preparations that could have gone to other uses (Fowler, 2016).

The median response for the error of commission group (taking action that was not needed) was five while that for the error of omission group (not taking action that was needed) was three—and remember, lower numbers indicate greater concern. Nearly 27% of the latter group indicated that they would be extremely concerned about the security of their job, while only 3% of the former group indicated this. Using a Somers' *D* test[5] the distribution of responses for the error of omission group is more skewed left, with lower numbers evidencing greater job security concern, than the error of commission group, at a 0.001 significance level.

We followed this question by asking both groups combined ($n = 133$) to indicate on a 7-point Likert-type scale their comfort with these errors relative to each other.[6] About 70% of these indicated that they would dislike an error of omission more than an error of commission (another 13% indicated they would equally dislike them, and only 17% that would dislike an error of commission more). The strength of dislike of an error is a form of regret about a decision made in the past.

The finding that emergency managers dislike errors of omission more than errors of commission runs contrary to findings in other studies and shows that decision biases and heuristics are relative to context. The bulk of previous experiments suggest that people have less regret or prefer to be wrong when the negative outcome is the result of an omission or failure to act than when the negative outcome is the result of an action they took, an error of commission (Baron & Ritov, 1994; Ordonez, Connolly & Coughlan, 2000; Connolly, Ordóñez, & Coughlan, 1997; Ritov & Baron, 1995). Most of the experimental evidence has been gathered through abstract games with random subject pools or college students, however, with only a few of the studies conducted in specific practical domains, including car insurance and vaccination decisions (D. A. Asch et al., 1994; Gilovich & Medvec, 1994; Johnson, Hershey, Meszaros, & Kunreuther, 1993). In all of these studies, people experienced less regret in errors of omission than in errors of commission.

Emergency managers are different because they are charged with preparing for potentially destructive events. When the forecast rain never falls, or the county is prepared and rain does not become a severe flood, the emergency manager is just doing her job. But when a massive flood destroys property and catches a county unawares, people point fingers and wonder why the emergency manager did not act sooner or more fully. These same blame dynamics may not hold in other public management fields where failure to act may be mere red tape with fewer discernable consequences, and taking action may mean sticking one's neck out and claiming responsibility for something for which one could be blamed later. Long (1949) famously described the bureaucrat's natural state as one of caution, but in emergency management and other fields such as financial accounting

that require planning for potential crisis, not taking action may be equally risky.

Attribution Bias Hypothesis

We see the possible presence of attribution bias when we consider two other groups of respondents to whom we presented our forecast experiment. These comprise 131 respondents who we asked to identify the same impacts of errors of commission and errors of omission as with the 133 respondents just discussed, but with respect to other emergency managers rather than themselves; that is, the language in our forecast experiment for these 131 respondents was about "an emergency manager who is unknown to you . . . who works in a different county."

The difference in the relative dislike of errors of omission and commission increased in this latter group of 131 respondents. When thinking of the impact of these errors on other emergency managers, 84% indicated that they thought other emergency managers would dislike an error of omission more than one of commission (as compared to 70% for the respondents reporting their own dislike). The difference between these two is statistically significant at the 0.01 level using a test of proportions. This experiment also shows the effects of framing on how an emergency manager evaluates a situation.

Numeracy Hypothesis

To test our numeracy hypothesis, we divided our respondents into four groups and presented each with a different version of an experiment about how they would respond to a long-term forecast predicting a destructive flood if action to protect the community were not taken right now. These four groups represented a two-by-two treatment of a highly likely versus less likely flood, and information expressed in frequencies versus probabilities. For instance, one group received a forecast expressed as 1 out of every 10 ($n = 61$), another a 0.1 probability ($n = 60$), another as 9 out of every 10 ($n = 61$), and the fourth a 0.9 probability ($n = 61$). Each respondent saw one scenario and answered two questions, "how acceptable to you is this probability of loss if action isn't taken?" and "how likely would you be to activate a costly emergency management plan . . ."

Our results showed that despite routinely encountering forecast information in their work, EMs did not respond the same way to numerically equivalent probabilities. Losses were deemed more acceptable—the scale ranged from 1 ("not at all acceptable") to 7 ("very acceptable")—when forecasts were expressed as a percentage rather than as a ratio, and EMs were less likely to activate an emergency management plan in the forecast of 0.1% than in the 1 out of 10 scenario. More than 30% of EMs given the "1 out of 10" uncertainty language said that the losses would be not at all acceptable, but less than 15% of EMs given the "0.1 probability" said so, although the outcomes are numerically equivalent. We also presented the same situation using the 9 out of 10 language and 0.9 probability. In this latter, higher likelihood pair, a higher percentage of EMs confronted with the "9 out of 10" presentation also found greater unacceptability of losses than did EMs confronted with the "0.9" probability language (41% vs. 19.7%). Both differences are statistically significant at the 0.05 level using the Somers' *D* test.[7]

For the "how likely would you be to activate a costly emergency management plan" question, responses extend from "not at all likely" (1) to activate the plan to "very likely" (7) to activate the plan. The high-likelihood (90%) forecasts not surprisingly elicited a higher likelihood of activating a plan than the low-likelihood (10%) ones, for both uncertainty formats. More revealingly, a much higher percentage of EMs given the 1 out of 10 language (68%) was likely to activate the plan (response of 5, 6, or 7 on a Likert-type scale) than those receiving the 0.1 language (43%). The differences in responses to the two language presentations are statistically significant at a 0.001 level. The differences are less pronounced when respondents were presented with the 90% forecast probability as either 0.9 or 9 in 10, and are significant at only a 0.1 level. While the 90% forecast probability is an interesting empirical example at the high end, managers are less likely to encounter 90% probability forecasts in their experience than lower-likelihood forecasts. Therefore, managers' reaction to a 10% forecast probability may reveal more about how managers approach their daily work than their reaction to an unusually high forecast probability.

Neighbor Effect Hypothesis

We asked managers to indicate the likelihood of implementing a costly emergency management plan in their jurisdiction in response to a forecast of a high probability flood, and then varied information on what neighboring counties were doing, to test our neighbor effect hypothesis. We found that a far lower proportion of the managers indicated that were "very likely" to implement the plan in their jurisdiction when they heard that neighboring counties had decided not to do so (14%) than when they heard the neighboring counties had decided to implement the plan (51%), although the actual forecast information was identical in the two situations. A similar gap appeared for a low-probability flood, and the differences between responses for each comparison were statistically significant at 0.001 level using a Somers' *D* test.

As discussed earlier, this could reflect elements of rational response to new or private information (what the neighbor did) and/or a mechanism to cope with uncertainty and anxiety about threats to performance and reputations. As with financial professionals whose buy and sell recommendations have an effect on their colleagues' recommendations (Welch, 2000), emergency management professionals likely

take the judgment of their neighboring colleagues seriously. This enables them to make more sense of complex, uncertain scientific information about flood forecasts through processing of the information via peer networks rather than as individual consumers of the information. Therefore, the neighbor effect shows that some decision biases and heuristics may be used to positive effect.

The presence of an effect provides useful information for changing practice. For example, county-level emergency managers have some discretion in deciding whether to adopt and implement certain kinds of plans to prepare for disasters that research shows are beneficial. If we observe a neighbor effect, then we would recommend that statewide planning efforts highlight managers that have already adopted hazard plans to persuade holdouts.

Outcome Bias Hypothesis

We presented 45 of our respondents with a scenario in which a group of emergency managers from three neighboring counties received a forecast of a flood and used a historical analog of the forecasted flood to decide to activate an emergency management plan. We also specified that the predicted flood happened but flood-related damages in all three counties were minimal. We then asked respondents for their opinion of the "quality of the group's decision-making process" as well as the effect they thought the group's decision would have on each member's reputation in their home county. We found that 98% indicated that the quality of the decision process was better than average, and 91% indicated that the reputation of members would be slightly to greatly enhanced by the group decision.

We then asked the same questions of a different group of 45 respondents (with no overlap between the two groups). With this second group of respondents, however, we specified that the predicted flood happened and very bad damages occurred. In this second group, 80% indicated that the quality of the decision process was better than average, and 58% indicated that the reputation of members would be slightly to greatly enhanced by the group decision. There are two particularly striking differences between the treatment responses.

First, notwithstanding the general similarity between the answers of these two groups of respondents, the responses show a statistically significant difference. For the question related to the quality of the decision process, the difference is significant at a 0.01 level, with the group with heavy flood damages providing lower rankings of decision quality than the group with minimal flood damages. For the reputation question, the difference is significant at the 0.001 level, with the heavy flood damage group indicating lower (less positive) effects on reputation. This implies that the end outcome of a flood event—minimal or heavy damages—influences perceptions about the quality of a decision-making process and the effects of the process on the professional reputations of the decision makers.

The presence of this outcome bias suggests that a process that evaluates information and takes action to prevent a harmful outcome will lead to more negative perceptions of decision-making quality, and of the decision makers themselves, if the harm ends up occurring than if it does not occur, even when the same evaluation process and action is followed in both cases. Similarly, emergency managers who performed well but were overwhelmed by the severity of a storm or by the lack of preparation by people and groups outside of their control also appear more likely to have their reputations suffer and have their plans rewritten. Rather than rely on outcome bias to shape planning for the next storm, even professional emergency managers need to engage in a deliberate process to write an after-action report and connect specific actions to threats recognizing that the next storm will not necessarily be the same as the last.

Implications

Our study represents the first test of prospect theory and decision biases and heuristics among emergency managers and the most comprehensive effort of the few studies done on biases and heuristics of public managers in their area of expertise. Its findings largely support each of our six hypotheses that emergency managers exhibit some of the same biases and heuristics found in the general public and college students, even in their domain of professional expertise. Table 3 highlights a subset of this evidence.

We build on an emerging tradition in behavioral public administration, a subfield that uses theories from psychology to explain administrative behavior at the individual level (Grimmelikhuijsen et al., 2017). Some behavioral work has been applied to understanding how cognitive biases shape work routines, such as the role of anchoring effects on performance appraisals (Belle, Cantarelli, & Belardinelli, 2017), or how a bias in favor of technology shapes decisions to acquire e-voting machines (Moynihan & Lavertu, 2012). Rather than administrative processes generally, however, we focus on county-level emergency managers deciding on options related to their professional expertise. Our project also contributes to public administration research that finds that improving management practice is one way to curb disaster losses (Kettl, 2006).

Information provided to politicians sometimes bears only a weak relationship to policy (Geys & Sørensen, 2018). Our project identifies a variety of cognitive biases that shape how information becomes policy. While behavioral economics and psychology also are concerned with cognitive limitations in this realm, public administration more directly takes into account how environmental pressures and incentive structures contribute to suboptimal decisions. Although we study emergency managers, we suspect all public managers are subject to biases and framing effects. Future research should identify how framing effects shape bureaucrats' decisions specifically, not just those of citizens or politicians, and how the implications of framing effects vary across fields (Olsen, 2015). Studying the microfoundations of

Table 3. Heuristics and Biases among Emergency Managers.

Description	Illustrative example
Prospect Theory	For a forecasted flood threatening 100 homes, 89% of respondents chose a risky action when the impacts were presented as losing houses. For an equally damaging flood with the action presented as saving houses, only 42% chose a risky action.
Regret: Regretting taking an action under uncertainty that in hindsight would have been beneficial (omission) or regretting taking an action under uncertainty that in hindsight would have been beneficial to avoid (commission).	Around 27% of respondents said they would be extremely concerned about the security of their job if they failed to take action in response to a flood forecast flood and the flood did occur, while only 3% of respondents said they would be extremely concerned about the security of their job if they took action in response to a flood forecast and the flood did not occur.
Attribution Bias: Assessing the decisions of others differently than one's own decision, often ascribing others' decision to personality and one's own decision to situational context.	Totally, 70% of respondents reported they would dislike an error of omission more than an error of commission if they made the error, but 84% of respondents reported that they thought other emergency managers would like an error of omission more than an error of commission.
Numeracy: Difficulties with processing numerical information, particularly probabilities compared to frequencies.	Totally, 30% of respondents indicated that they would likely activate an emergency management plan in response to a flood forecast with a one in 10 chance of happening, while only 22% of respondents indicated that they would likely activate an emergency management plan in response to a flood forecast with a 10% chance of happening.
Neighbor Effect: The influence on behavior of peers' behavior, such that one goes along with actions that others have taken without normal consideration of the consequences and desired results.	About 51% of respondents said they would likely take action in response to a particular forecast when they learned other counties were taking action, and only 14% of respondents said they would likely take action in response to an identical forecast when they learned other counties were not taking action.
Outcome Bias: People rate decision quality as better, or the decision maker as more competent, when the outcome was favorable.	Around 98% said that the quality of the decision process was better than average, and 91% indicated that the reputation of members would be enhanced when the group enacted a plan and a flood happened with minimal damages. However, when very bad damages occurred, 80% indicated that the quality of the decision process was better than average, and 58% indicated that the reputation of members would be enhanced by the group decision.

administrative behavior holds promise for advancing public administration more broadly through one path identified in this article (Moynihan, 2018), examining how biases and framing effects shape decisions of public managers in their professional fields.

While showing that the biases found among college students and the general public appear present among professionals making decisions in their areas of expertise, we have ignored the question of why this is the case? One explanation may be simply that emergency managers are not more strategic or rational than other managers or the general public (Loewen, Sheffer, Soroka, Walgrave, & Shaefer, 2014). Expertise may not operate as a general moderating factor improving decision quality. There is a long-standing debate in social science between the rational choice school, which assumes that elites respond to incentives and act strategically, and behaviorists from psychology and economics who assume that behavioral anomalies result from humans' cognitive architecture and therefore are subject to manipulation (Riccucci, 2010, pp. 21-64). Rational choice scholars often study politicians and political elites, while behavioralists study voters. We provide evidence that emergency managers correspond to the behavioralists' portrayal, although the debate whether public administrators exhibit strategic behavior more like politicians or behavioral anomalies more like voters is far from over.

The presence of decision biases naturally leads to an investigation into what to do about them. One remedy to individual-level departures from rationality may be structured decision processes at the organizational level. Structured decision-making processes divide a problem into stages and use facilitators to allow participants to more explicitly define objectives, detail performance metrics, construct alternative courses of action, and face tradeoffs. Empirical studies of structured decision making (SDM) show promise for mitigating some of both individual and group-level decision constraints, with most studies being applied to the environmental resource management context (Arvai & Gregory, 2003; Gregory et al., 2012; Gregory & Long, 2009). Some might say that structured decision

processes are an obvious way forward, but in practice they are only obvious in retrospect after a failure as most decision processes proceed through rote procedure, intuitive guesswork, or muddling through.

The presence of decision biases among emergency managers also has implications for the use of behavioral science by government to frame decisions for the public (Lodge & Wegrich, 2016; Oliver, 2015). If government managers themselves are subject to biases and framing effects, then decision processes should attempt to mitigate these effects and incorporate relevant information about the decision context rather than assume that government managers operate from a position of objectivity. This is not an argument against the use of the "nudge," as no position on a policy issue is entirely neutral, but instead a call to consider the manager's decision context and subject her decisions to analysis from multiple perspectives (Sunstein, 2016, pp. 18-20). Having said this, it is a fair question to ask whether our study captures real-world decision dynamics well enough to make such recommendations for practice.

Our study partially addresses questions about external validity by posing questions to professional emergency managers about their domain of expertise, rather than relying on proxies such as college students or a random sample of the public about issues that they do not regularly confront. We also have employed these experiments face-to-face in professional conferences of emergency managers and found in debriefings that these managers find the decision context real and relevant. To address internal validity, our study employs a survey experiment with individuals randomly assigned to different treatments. We pose questions that our respondents face in their work and are competent to understand. Unfortunately, our sample size and research design that splits this sample into four groups makes it difficult to examine with power and confidence the effects of age, gender, experience, and other demographic and professional characteristics on biases.[8] We not only acknowledge this limitation but also offer it as an argument for future research to examine the import of such differences on decision making.[9]

Finally, we anticipate a possible criticism that the findings may not travel to other public management domains. However, we built our study around hypotheses drawn from prospect theory and psychological studies of biases and heuristics that have found support in other fields outside of public management. Emergency managers are likely more exposed to decision making under uncertainty than other managers, and thus serve as a canary in the mine. If they are subject to decision biases and employ heuristics in the face of risk and uncertainty although they routinely face these problems, then other managers may do the same. Understanding and improving decision processes should be a priority for public management more broadly.

Moving to recommendations, if emergency managers are more risk averse when the outcomes of actions are framed as gains than when equivalent outcomes are framed as losses—a finding consistent with prospect theory—then they should receive information in multiple forms (Abdellaoui et al., 2007; Kahneman & Tversky, 1979). Future research can explore whether managers would be more likely to go for broke in planning to protect a town from flooding when the decision is framed in loss terms—and thus depleting the treasury—than when the decision is framed as a gain. Politicians and city managers should also be aware that their decision frames could affect such managers' calculations.

The results related to numeracy challenges—the ability of emergency managers to process numerical information consistently—are mixed, and suggest that no one form of information is sufficient to communicate a fact. Statistically significant differences in the acceptability of losses when forecasts are expressed in percentage rather than ratio forms, as well as significant differences in emergency managers' responses to a question about activating a costly emergency management plan depending on how the forecast likelihood is presented, suggest that percentage or ratio terms alone are insufficient in a forecast. State and national weather services offices, for example, can provide multiple forms of forecast information to give managers a complete picture, rather than assuming that mathematically equivalent presentations of information are understood as equivalent.

The presence of a neighbor effect suggests that there may be potential for pressure to reach premature consensus in group decision-making settings. By relying on the neighbor effect heuristic and being subject to pressures to reach consensus, managers may fail to create enough decisions alternatives. In contrast, the evidence suggests that group settings where diverse stakeholders more readily can express contradictory perspectives are better able to identify creative alternatives and win-win solutions than groups under time pressure, or those where opinions are squelched (Bryson, Quick, Slotterback, & Crosby, 2013). We encourage emergency management decision processes to bring together multiple stakeholders and make decision alternatives explicit when important issues are at stake so as not to rely solely on neighbor effects. At the same time, a neighbor effect could be used as a nudge to encourage emergency managers to take action. For example, if a manager knows that another manager in a neighboring county has updated her emergency contact list—with outdated contact lists being a common problem during a disaster—then he may be more likely to do so.

Finally, the presence of attribution bias suggests that emergency managers may devalue the context-specific information that shapes decision making in other environments, and the diversity of experiences that other managers bring. A failure to recognize context specificity and diverse backgrounds can be an impediment to group decision making. The presence of attribution bias among our sample is merely suggestive of such problems in group decision making, and the topic deserves further study. Managers could

reach a false consensus, or the biases and heuristics could exacerbate tendencies for individuals to remain immune to constructive negative feedback.

The next step is to further study how public managers in a variety of contexts approach decision making under conditions of risk and uncertainty. Most of our respondents (~90%) agree that they would benefit from additional training and support when it comes to making complex decisions under uncertainty. Given evidence that emergency managers use heuristics and are susceptible to biases that may have negative consequences for decision making, scholars and managers should increase the use of more formal decision support processes, such as presenting information in public forms, or employing processes that clarify norms and values, rank alternatives, and identify tradeoffs (Gregory et al., 2012). The evidence suggests that nudges could be used on public managers—if an appropriate behavioral target such as updating contact lists can be identified. At the very least, concrete proposals for public managers to nudge citizens should not assume that the managers operate from a neutral position of objective rationality. Theories from psychology and behavioral economics have been employed to study how public opinion departs from rational utility models. It is time for the study of bureaucrats' own departures from the rational model to inform how to structure government's interactions with the public.[10]

Acknowledgments

The authors would like to acknowledge the advice and assistance of Joseph Arvai and Kelly Redmond.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project was funded by National Science Foundation Award 1133263, DRRC/Collaborative Research: Emergency Management in Rural America: Decision-Makers Use of Climate Science in Flood Planning and Management.

Notes

1. Rural counties comprise 42.5% of all counties in the United States. We use the 13% higher survey response rate of local metro health departments relative to local nonmetro health departments reported in Timsina (2017) to stratify our sample frame to include a roughly 48% share of rural counties.
2. Definition available at: https://www.census.gov/programs-surveys/metro-micro.html, accessed 7/19/2018.
3. Thus, for example, 65 emergency managers belong in Group C in our second set of experiments and 65 belong in Group C in our third set of experiments, but only 18 of the 65 in each group are common to both groups. This treatment design obviates possible confounding from the order of questions, as each respondent sees only one treatment in each of our decision heuristic categories. Because our survey software randomly assigns each emergency manager to each experimental set independently of the other experimental sets, our group sample sizes also vary slightly within each set. Higher nonresponse rates for some questions (e.g., our outcome bias experiment) generated substantial group size differences between experimental sets, although the group sizes within any given set remain comparable. Finally, in most of the analyses that we report below we compare only two (out of four) groups at a time, yielding group sizes of 60 to 80 emergency managers depending on the question. However, some of the questions in our attribution bias experiment have only two versions, one seen by two of the groups and the other by the other two groups. For those questions, our subsamples will be twice the typical size (roughly 130 respondents).
4. We performed a permutation test using the randomization inference "ritest" routine in Stata for this and all following statistical analyses that we report to examine whether our results appear sensitive to the random assignment of respondents to questionnaire treatments discussed earlier (for more details, see Heß, 2017). For each analysis, we generated 1,000 replications. For two of our results with significance at the 0.05 level, the upper bound of the p value confidence interval exceeded 0.05. We indicate these exceptions in a footnote to the relevant results. All of our other results appear robust to this test, suggesting that they are not artifacts of our assignment process.
5. We use Somers' D for ordinal scale responses to examine the statistical significance of differences in between groups responding to our different scenarios. Somers' D, a nonparametric test (Newson, 2002), provides a formal test of the association between our binary variable representing the scenario treatment and our ordinal responses.
6. Following the presentation of forecast information about a flood risk, the text of this question read as follows:
   Consider the following two possible outcomes that could be realized in light of decisions you make:
   Outcome 1: You take action to minimize the flood risk, and the expected flood DOES NOT occur.
   Outcome 2: You do not take action to minimize the flood risk and the expected flood DOES occur.
   Circle the number that corresponds best with your opinion.
   Choices ranged over a 7-point scale, with one end labeled "I would STRONGLY dislike Outcome 1 more than Outcome 2," the middle choice labeled "I would dislike Outcome 1 and Outcome 2 EQUALLY," and the other end labeled "I would STRONGLY dislike Outcome 2 more than Outcome 1."
7. Our permutation test yields a p value upper bound of 0.075 and 0.07 for the low- and high-likelihood forecast, respectively (see footnote 4).
8. Following the suggestion of a reviewer, we reran our analyses and included only the most experienced emergency managers in our sample, those with more than 15 years of professional experience. The results follow those reported above for the full sample for most of the biases. The exceptions are (a) weaker significance at the 0.1 level for the attribution bias test (similar difference in means between the treatments but lower significance due to smaller sample size), (b) insignificance in the numeracy bias test for differences in the

acceptability of losses in the numeracy experiment with the 0.1 likelihood forecast (but the numeracy bias test for differences in the likelihood of activating the emergency plan are equally significant as in the full sample), and (c) insignificance for both the acceptability of losses and likelihood of activating the emergency plan in the numeracy experiment for the 0.9 likelihood forecast. Those who spend more than 80% of their time working on flood issues show nearly identical results.

9. We also recognize that the statistical significance of the 11 associations that we presented independently fails to account for the effect of multiple comparisons (Althouse, 2016; Ranganathan, Pramesh & Buyse, 2016); that is, the likelihood of making at least one Type I error, an incorrectly rejected null hypothesis of no association, increases as we increase the number of tests. In fact, we do not provide *p* values for each test partly because they become less meaningful in such a multiple comparison context. To address this potential concern, we conducted family-wise error rate step-down (Holland & Copenhaver, 1987) and "false discovery rate" step-up and step-down rate (Benjamini & Liu, 1999; Simes, 1986) corrections to the joint probability statements. Two of these (the false discovery corrections) maintained joint significance at the above-stated levels for all associations, while the more conservative Holland correction failed to reject at the 0.05 level the null hypotheses for the two numeracy tests associated with the acceptability of losses (rejecting these only at the 0.1 level).
10. National Science Foundation award number 1133024, National Oceanic and Atmospheric Administration award number NA13OAR4310116, and the Institute for Society, Culture and Environment at Virginia Tech partially supported the work reported here. We also would like to acknowledge the contributions of Joseph Arvai and the late Kelly Redmond to our thinking. However, this article represents our views and not necessarily those of the many who have provided advice or financial support.

ORCID iD


Patrick S. Roberts https://orcid.org/0000-0002-8920-0396

References

Abdellaoui, M., Bleichrodt, H., & Paraschiv, C. (2007). Loss aversion under prospect theory: A parameter-free measurement. *Management Science*, *53*, 1659-1674.

Althouse, A. D. (2016). Adjust for multiple comparisons? It's not that simple. *The Annals of Thoracic Surgery*, *101*, 1644-1645.

American Association for Public Opinion Research. (2016). *Standard definitions: Final dispositions of case codes and outcome rates for surveys*. Oakbrook Terrace, IL: Author.

Arvai, J., & Gregory, R. (2003). Testing alternative decision approaches for identifying cleanup priorities at contaminated sites. *Environmental Science & Technology*, *37*, 1469-1476.

Arvai, J., Gregory, R., Ohlson, D., Blackwell, B., & Gray, R. (2006). Letdowns, wake-up calls, and constructed preferences: People's responses to fuel and wildfire risks. *Journal of Forestry*, *104*, 173-181.

Asch, D. A., Baron, J., Hershey, J. C., Kunreuther, H., Meszaros, J., Ritov, I., & Spranca, M. (1994). Omission bias and pertussis vaccination. *Medical Decision Making*, *14*, 118-123.

Asch, S. E. (1956). Studies of independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs: General and Applied*, *70*(9), 1-70.

Baekgaard, M. (2017). Prospect theory and public service outcomes: Examining risk preferences in relation to public sector reforms. *Public Administration*, *95*, 927-942.

Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, *54*, 569-579.

Baron, J., & Ritov, I. (1994). Reference points and omission bias. *Organizational Behavior and Human Decision Processes*, *59*, 475-498.

Belle, N., Cantarelli, P., & Belardinelli, P. (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*, 275-294.

Bendor, J., Taylor, S., & Van Gaalen, R. (1985). Bureaucratic expertise versus legislative authority: A model of deception and monitoring in budgeting. *American Political Science Review*, *79*, 1041-1060.

Benjamini, Y., & Liu, W. (1999). A step-down multiple hypotheses testing procedure that controls the false discovery rate under independence. *Journal of Statistical Planning and Inference*, *82*, 163-170.

Bryson, J. M., Quick, K. S., Slotterback, C. S., & Crosby, B. C. (2013). Designing public participation processes. *Public Administration Review*, *73*, 23-34.

Chen, Y. F. (2008). Herd behavior in purchasing books online. *Computers in Human Behavior*, *24*, 1977-1992.

Christie, W. G., & Huang, R. D. (1995). Following the pied piper: Do individual returns herd around the market? *Financial Analysts Journal*, *51*, 31-37.

Comfort, L. K., Waugh, W. L., & Cigler, B. A. (2012). Emergency management research and practice in public administration: Emergence, evolution, expansion, and future directions. *Public Administration Review*, *72*, 539-547.

Connolly, T., Ordóñez, L. D., & Coughlan, R. (1997). Regret and responsibility in the evaluation of decision outcomes. *Organizational Behavior and Human Decision Processes*, *70*, 73-85.

Cooper, J. C., & Kovacic, W. E. (2012). Behavioral economics: Implications for regulatory behavior. *Journal of Regulatory Economics*, *41*, 41-58.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, *58*, 1-73.

Devenow, A., & Welch, I. (1996). Rational herding in financial economics. *European Economic Review*, *40*, 603-615.

Druckman, J. N., & McDermott, R. (2008). Emotion and the framing of risky choice. *Political Behavior*, *30*, 297-321.

Eiser, J. R., Bostrom, A., Burton, I., Johnston, D. M., McClure, J., Paton, D., . . . White, M. P. (2012). Risk interpretation and action: A conceptual framework for responses to natural hazards. *International Journal of Disaster Risk Reduction*, *1*, 5-16.

Ericsson, K. A., Hoffman, R. R., Kozbelt, A., & Williams, A. M. (Eds.). (2018). *The Cambridge handbook of expertise and expert performance*. New York: Cambridge University Press.

Evans, S. C., Roberts, M. C., Keeley, J. W., Blossom, J. B., Amaro, C. M., Garcia, A. M., . . . Reed, G. M. (2015). Vignette methodologies for studying clinicians' decision-making: Validity, utility, and application in ICD-11 field studies. *International Journal of Clinical and Health Psychology*, *15*, 160-170.

Fatas, E., Neugebauer, T., & Tamborero, P. (2007). How politicians make decisions: A political choice experiment. *Journal of Economics*, *92*, 167-196.

Feeney, M. K., & Smith, C. R. (2008). Social embeddedness in outsourcing: What shapes public managers' perceptions? *Public Performance & Management Review*, *31*, 517-546.

Fowler, S. (2016, September 8). Madison Co. emergency management director fired. *Clarion-Ledger*. Retrieved from https://www.clarionledger.com/story/news/2016/09/08/madison-emergency-management-director-fired/89996192/

Geys, B., & Sørensen, R. J. (2018). Never change a winning policy? Public sector performance and politicians' preferences for reforms. *Public Administration Review*, *78*, 206-216.

Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology (and vice versa). In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 129-161). Oxford, England: John Wiley & Sons.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102*, 684-704.

Gilovich, T., & Medvec, V. H. (1994). The temporal pattern to the experience of regret. *Journal of Personality and Social Psychology*, *67*, 357-365.

Gill, J., & Meier, K. J. (2000). Public administration research and practice: A methodological manifesto. *Journal of Public Administration Research and Theory*, *10*, 157-199.

Giwargis, R. (2017, March 6). San Jose flood: City hires new emergency services director. *San Jose Mercury News*. Retrieved from https://www.mercurynews.com/2017/03/06/san-jose-flood-city-hires-new-emergency-services-director/

Gregory, R., Failing, L., Harstone, M., Long, G., McDaniels, T., & Ohlson, D. (2012). *Structured decision making: A practical guide to environmental management choices*. Hoboken, NJ: John Wiley.

Gregory, R., & Long, G. (2009). Using structured decision making to help implement a precautionary approach to endangered species management. *Risk Analysis: An International Journal*, *29*, 518-532.

Grimmelikhuijsen, S., Jilke, S., Olsen, A. L., & Tummers, L. (2017). Behavioral public administration: Combining insights from public administration and psychology. *Public Administration Review*, *77*, 45-56.

Hafner-Burton, E. M., Hughes, D. A., & Victor, D. G. (2013). The cognitive revolution and the political psychology of elite decision making. *Perspectives on Politics*, *11*, 368-386.

Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences*, *112*, 2395-2400.

Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*, *107*, 311-327.

Heß, S. (2017). Randomization inference with Stata: A guide and software. *Stata Journal*, *17*, 630-651.

Heider, F. (1958). *The psychology of interpersonal relations*. New York, NY: John Wiley.

Heimann, C. L. (1993). Understanding the Challenger disaster: Organizational structure and the design of reliable systems. *American Political Science Review*, *87*, 421-435.

Holland, B. S., & Copenhaver, M. D. (1987). An improved sequentially rejective Bonferroni test procedure. *Biometrics*, *43*, 417-423.

International Association of Emergency Managers. (2005). *EM demographics survey—2005* (Compiled by C. Marks). Retrieved from https://training.fema.gov/hiedu/surveys.aspx

Janis, I. L. (1982). *Groupthink: Psychological studies of policy decisions and fiasces*. Boston, MA: Houghton Mifflin.

Jensen, J., & Youngs, G. (2015). Explaining implementation behaviour of the national incident management system (NIMS). *Disasters*, *39*, 362-388.

Johnson, E. J., Hershey, J., Meszaros, J., & Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, *7*, 35-51.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*, *5*, 193-206.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, *47*, 263-291.

Kahneman, D., & Tversky, A. (1984). Choices, values and frames. *American Psychologist*, *39*, 342-350.

Kastanakis, M. N., & Balabanis, G. (2012). Between the mass and the class: Antecedents of the "bandwagon" luxury consumption behavior. *Journal of Business Research*, *65*, 1399-1407.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation* (pp. 192-223). Lincoln: University of Nebraska Press.

Kettl, D. F. (2006). Is the worst yet to come? *The Annals of the American Academy of Political and Social Science*, *604*, 273-287.

Kinder, D. R., & Palfrey, T. R. (Eds.). (1993). *Experimental foundations of political science*. Ann Arbor, USA: University of Michigan Press.

Klein, G. (2008). Naturalistic decision making. *Human Factors*, *50*, 456-460.

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

Kreps, D. M., & Wilson, R. (1982). Reputation and imperfect information. *Journal of Economic Theory*, *27*, 253-279.

Krueger, J., & Clement, R. W. (1994). The truly false consensus effect: An ineradicable and egocentric bias in social perception. *Journal of Personality and Social Psychology*, *67*, 596-610.

Landau, M., & Stout, R. (1979). To Manage Is Not to Control: Or the Folly of Type II Errors. *Public Administration Review*, *39*(2), 148-156. doi:10.2307/3110470

Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes*, *76*, 149-188.

Lodge, M., & Wegrich, K. (2016). The rationality paradox of nudge: Rational tools of government in a world of bounded rationality. *Law & Policy*, *38*, 250-267.

Loewen, P., Sheffer, L., Soroka, S., Walgrave, S., & Shaefer, T. (2014, October 6). *Expertise and efficacy in elite political decision making* (Working paper). Retrieved from https://sites.duke.edu/2014bmp/files/2014/10/Loewen_et_al.pdf

Long, N. (1949). Power and administration. *Public Administration Review*, *9*, 257-264.

McDermott, R., Fowler, J. H., & Smirnov, O. (2008). On the evolutionary origin of prospect theory preferences. *The Journal of Politics*, *70*, 335-350.

Mercer, J. (2005). Prospect theory and political science. *Annual Review of Political Science*, *8*, 1-21.

Mieg, H. A. (2006). Social and sociological factors in the development of expertise. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (743-760). New York: Cambridge university Press.

Morris, J. C., Morris, E. D., & Jones, D. M. (2007). Reaching for the philosopher's stone: Contingent coordination and the military's response to Hurricane Katrina. *Public Administration Review*, *67*, 94-106.

Moynihan, D. P. (2008). Learning under uncertainty: Networks in crisis management. *Public Administration Review*, *68*, 350-365.

Moynihan, D. P. (2018). A great schism approaching? Towards a micro and macro public administration. *Journal of Behavioral Public Administration*, *1*, 1-8.

Moynihan, D. P., & Lavertu, S. (2012). Cognitive biases in governing: Technology preferences in election administration. *Public Administration Review*, *72*, 68-77.

Murphy, A. H., & Winkler, R. L. (1977). Can weather forecasters formulate reliable probability forecasts of precipitation and temperature. *National Weather Digest*, *2*(2), 2-9.

Newson, R. (2002). Parameters behind "nonparametric" statistics: Kendall's tau, Somers' D and median differences. *Stata Journal* 2 (1), 45-64.

Nuclear Regulatory Commission. (1986). *Loss of integrated control system power and overcooling transient at Rancho Seco on December 26, 1985* (No. NUREG–1195). Washington, DC: Author.

Oliver, A. (2015). Nudging, shoving, and budging: Behavioural economic-informed policy. *Public Administration*, *93*, 700-714.

Olsen, A. L. (2015). Citizen (dis)satisfaction: An experimental equivalence framing study. *Public Administration Review*, *75*, 469-478.

Ordóñez, L. D., Connolly, T., & Coughlan, R. (2000). Multiple reference points in satisfaction and fairness assessment. *Journal of Behavioral Decision Making*, *13*, 329-344.

Payne, P., & Morris, J. D. (2018). "Sonoma county emergency official reassigned amid public criticism over fire alerts." The Press Democrat. Sonoma, CA. February 21.

Peabody, J. W., Luck, J., Glassman, P., Jain, S., Hansen, J., Spell, M., & Lee, M. (2004). Measuring the quality of physician practice by using clinical vignettes: A prospective validation study. *Annals of Internal Medicine*, *141*, 771-780.

Peerbolte, S. L., & Collins, M. L. (2013). Disaster management and the critical thinking skills of local emergency managers: Correlations with age, gender, education, and years in occupation. *Disasters*, *37*, 48-60.

Peters, E., Västfjäll, D., Gärling, T., & Slovic, P. (2006). Affect and decision making: A "hot" topic. *Journal of Behavioral Decision Making*, *19*, 79-85.

Plott, C. R. (1982). Industrial organization theory and experimental economics. *Journal of Economic Literature*, *20*, 1485-1527.

Plott, C. R., & Zeiler, K. (2007). Exchange asymmetries incorrectly interpreted as evidence of endowment effect theory and prospect theory? *American Economic Review*, *97*, 1449-1466.

Rahm, D., & Reddick, C. G. (2011). US city managers' perceptions of disaster risks: Consequences for urban emergency management. *Journal of Contingencies and Crisis Management*, *19*, 136-146.

Ranganathan, P., Pramesh, C. S., & Buyse, M. (2016). Common pitfalls in statistical analysis: The perils of multiple testing. *Perspectives in Clinical Research*, *7*, 106-107.

Reyna, V. F., & Brainerd, C. J. (2008). Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences*, *18*, 89-107.

Riccucci, N. M. (2010). *Public administration: Traditions of inquiry and philosophies of knowledge*. Washington, USA: Georgetown University Press.

Ritov, I., & Baron, J. (1995). Outcome knowledge, regret, and omission bias. *Organizational Behavior and Human Decision Processes*, *64*, 119-127.

Roberts, P. S., & Wernstedt, K. (2016). Using climate forecasts across a state's emergency management network. *Natural Hazards Review*, *17*(3), 05016002-1- 05016002-11.

Rook, L. (2006). An economic psychological approach to herd behavior. *Journal of Economic Issues*, *40*, 75-95.

Rose, N. (1993). Government, authority and expertise in advanced liberalism. *Economy and Society*, *22*, 283-299.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed) *Advances in experimental social psychology* (pp. 173-220). New York, NY: Academic Press.

Schnellenbach, J., & Schubert, C. (2014). *Behavioral public choice: A survey* (Freiburger Diskussionspapiere zur Ordnungsökonomik, No. 14/03). Freiburg, Germany: University of Freiburg.

Shiller, R. J. (1995). Conversation, information, and herd behavior. *The American Economic Review*, *85*, 181-185.

Simes, R. J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika*, *73*, 751-754.

Simon, H. A. (1947). *Administrative behavior. A study of decision-making processes in administrative organization*. New York, NY: Free Press.

Slovic, P., Monahan, J., & MacGregor, D. G. (2000). Violence risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and Human Behavior*, *24*, 271-296.

Snary, C. (2004). Understanding risk: The planning officers' perspective. *Urban Studies*, *41*, 33-55.

Somin, I. (2017, October 13). Can we trust government to correct our cognitive biases? *The Washington Post*. Retrieved from https://www.washingtonpost.com/news/volokh-conspiracy/

wp/2017/10/13/can-we-trust-government-to-correct-our-cognitive-biases/?utm_term=.d9a10bf901ad

Stewart, T. R., Roebber, P. J., & Bosart, L. F. (1997). The importance of the task in analyzing expert judgment. *Organizational Behavior and Human Decision Processes*, *69*, 205-219.

Sunstein, C. R. (2016). *The ethics of influence: Government in the age of behavioral science*. New York: Cambridge University Press.

Swets, J. A. (1973). The relative operating characteristic in psychology: A technique for isolating effects of response bias finds wide use in the study of perception and cognition. *Science*, *182*, 990-1000.

Thaler, R. H., & Sunstein, C. R. (2003). Libertarian paternalism. *American Economic Review*, *93*, 175-179.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.

Timsina, L. R. (2017). Examining the activities, effectiveness, and contribution of local health departments using a national longitudinal survey of public health systems. *Theses and Dissertations-Epidemiology and Biostatistics*. 12. Retreived from https://uknowledge.uky.edu/epb_etds/12

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453-458.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297-323.

Weaver, J., Harkabus, L. C., Braun, J., Miller, S., Cox, R., Griffith, J., & Mazur, R. J. (2014). An overview of a demographic study of United States emergency managers. *Bulletin of the American Meteorological Society*, *95*, 199-203.

Welch, I. (2000). Herding among security analysts. *Journal of Financial Economics*, *58*, 369-396.

Wong, S. S., & Boh, W. F. (2010). Leveraging the ties of others to build a reputation for trustworthiness among peers. *Academy of Management Journal*, *53*, 129-148.

Wright, J. D. (2007). Behavioral law and economics, paternalism, and consumer contracts: An empirical perspective. *NYU Journal of Law & Liberty*, *2*(3). Retrieved from https://ssrn.com/abstract=1015899

Zaksek, M., & Arvai, J. L. (2004). Toward improved communication about wildland fire: Mental models research to identify information needs for natural resource management. *Risk Analysis: An International Journal*, *24*, 1503-1514.

Zhao, L., Yang, G., Wang, W., Chen, Y., Huang, J. P., Ohashi, H., & Stanley, H. E. (2011). Herd behavior in a complex adaptive system. *Proceedings of the National Academy of Sciences*, *13*, 15058-15063.

Author Biographies

Patrick S. Roberts is an associate professor in the Center for Public Administration and Policy, School of Public and International Affairs, at Virginia Tech in Alexandria, Virginia.

Kris Wernstedt is a professor in the School of Public and International Affairs at Virginia Tech in Alexandria, VA. He spent the last two years at the Institute of Human Settlements Studies at Ardhi University in Dar es Salaam, Tanzania.